

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 09:25:51 ; Search time 2621 Seconds

(without alignments)  
15034.414 Million cell updates/sec

Title: US-10-047-825-3

Perfect score: 1354  
Sequence: 1 aatcgagcagagcttaagc.....aaaaaaaaaaaaaaaaaaaaa 1354Scoring table: OLIGO\_MDC  
Gapop 60.0 ; Gapext 60.0

Searched: 2054640 segs, 14551402878 residues

Word size : 16

Total number of hits satisfying chosen parameters: 317414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pt:\*  
10: gb\_rc:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_ov:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_rc:\*  
27: em\_scs:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rtd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vtc:\*  
38: em\_sy:\*  
39: em\_hugo\_hum:\*  
40: em\_hugo\_mus:\*  
41: em\_hugo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	63	4.7 1335	CCU12314	U12314 Cenchrus ci
2	39	2.9 160562	AP005613	AP005613 Oryza sat
3	34	2.5 149417	AP003391	AP003391 Oryza sat
4	33	2.4 1496	BC005980	BC005980 Homo sapi
5	33	2.4 135792	AP003544	AP003544 Oryza sat
6	33	2.4 158456	AP004731	AP004731 Oryza sat
7	32	2.4 950	AF384374	AF384374 Lycopersi
8	32	2.4 186896	AC019057	AC019057 Homo sapi
9	31	2.3 1253	AK026389	AK026389 Homo sapi
10	31	2.3 1495	HS0804954	HS0804954 Homo sapi
11	31	2.3 1844	AX055708	AX055708 Sequence
12	31	2.3 1844	AX089940	AX089940 Sequence
13	31	2.3 1844	AX454590	AX454590 Sequence
14	31	2.3 1844	AX464542	AX464542 Sequence
15	31	2.3 1844	AX491068	AX491068 Sequence
16	31	2.3 3026	AX253072	AX253072 Sequence
17	31	2.3 5964	AX344763	AX344763 Sequence
18	31	2.3 6577	CE049263	CE049263 Caenorhabdi
19	31	2.3 56039	AC115598	AC115598 Dictyoste
20	31	2.3 106239	AL607122	AL607122 Human DNA
21	31	2.3 149599	AC112375	AC112375 Rattus no
22	31	2.3 152668	AC128193	AC128193 Rattus no
23	31	2.3 153798	AC108343	AC108343 Rattus no
24	30	2.2 121	E17121	E17121 DNA probe.
25	30	2.2 2086	AX146625	AX146625 Sequence
26	30	2.2 2704	AT130878	AT130878 Arabidops
27	30	2.2 3208	BC031926	BC031926 Homo sapi
28	30	2.2 14993	AY040832	AY040832 Hordelum v
29	30	2.2 50095	AC074833	AC074833 Mus muscu
30	30	2.2 64065	AC120003	AC120003 Mus muscu
31	30	2.2 65694	AC091065	AC091065 Homo sapi
32	30	2.2 109514	AC084722	AC084722 Homo sapi
33	30	2.2 135197	AL513319	AL513319 Homo sapi
34	30	2.2 137590	AC027567	AC027567 Homo sapi
35	30	2.2 140011	AL845337	AL845337 Mus muscu
36	30	2.2 143552	AC105021	AC105021 Homo sapi
37	30	2.2 144604	AC104974	AC104974 Homo sapi
38	30	2.2 155369	AC099166	AC099166 Rattus no
39	30	2.2 158641	AC102786	AC102786 Mus muscu
40	30	2.2 164347	AC034124	AC034124 Homo sapi
41	30	2.2 165921	AC019354	AC019354 Homo sapi
42	30	2.2 172350	AC079234	AC079234 Homo sapi
43	30	2.2 178737	AC015521	AC015521 Homo sapi
44	30	2.2 182286	AC025981	AC025981 Homo sapi
45	30	2.2 191654	AC022616	AC022616 Homo sapi

## ALIGNMENTS

RESULT 1  
CCU12314  
LOCUS Cenchrus ciliaris clone PX7 peroxidase mRNA, complete cds.  
DEFINITION U12314  
ACCESSION U12314  
VERSION U12314.1 GI:520567  
KEYWORDS  
SOURCE Cenchrus ciliaris.  
ORGANISM Cenchrus ciliaris.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Paniceae; Cenchrus.  
REFERENCE  
1 (bases 1 to 1335)  
AUTHORS  
TITLE  
Investigation of peroxidase genes and genetic transformation in

```

JOURNAL      bufel grass
REFERENCE     Thesis (1994) Botany, University of Queensland
AUTHORS      2 (bases 1 to 1335)
TITLE        Direct Submission
JOURNAL      Submitted (15-JUN-1994) Annette H. Ross, University of Queensland,
              Botany, St. Lucia, Brisbane, Queensland, 4072, Australia
FEATURES
SOURCE       1.1335
              /organism="Cenchrus ciliaris"
              /cultivar="Biloeia"
              /db_xref="taxon:35872"
              /clone="PX7"
              /tissue_type="wounded and unwounded leaves and stems"
              /clone_id="lambda gt-10"
              1. 57
              58. 981
              /BC_number="1.11.1.7"
              /note="expressed in leaves and is wound-induced"
              /codon_start=1
              /product="peroxidase"
              /protein_18="AAA20472.1"
              /db_xref="GI:520568"
              /translation="MAAPTLMOCLLAISLSTTANAQLSTFTVASSCPNIQIVYRAM
              TOAVSSEPRMGASILRFPHDFVQCDASILSRGERSAGPANVSRYEYIDIK
              KNEVACPGVYSCATIVPLARPGPNLIGPTWNPVILGRDSTAMLSLANOLPPPT
              SLGTLISLFGRLSARDMIALSGAHNAQRCCTFGRITGDTNIDASFALQOOTCPR
              SGGDNLPIDAQTPARFDNAYTYTNLVSRRGLFHSDDLEFNGGSODALVROYSSPSQ
              FNSDFVAMIMKGNIGANAGVRRNCRVVNS"
              58. 126
              /note="putative"
              127. 978
              /product="peroxidase"
              982. 1335
              1335
              3'UTR
              polyA_site
              /note="39 A residues"
              264 a 440 c 374 g 257 t
BASE COUNT   264 a 440 c 374 g 257 t
ORIGIN
Query Match      4.7%; Score 63; DB 8; Length 1335;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 826 TTCACCTCGACGACGAGCTTTCACGCGGCGGTCGACGAGACCGCTGTGAGCAGTAC 885
|||||
DB 808 TTCACCTCGACGACGAGCTTTCACGCGGCGGTCGACGAGACCGCTGTGAGCAGTAC 867
|||||
OY 886 AGC 888
|||||
DB 868 AGC 870

RESULT 2
AP005613/c 160562 bp DNA linear HTG 02-AUG-2002
LOCUS      Oryza sativa (japonica cultivar-group) chromosome 2 clone
DEFINITION OSJNBA0082C09, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION  AP005613
VERSION     AP005613.1 GI:22091085
KEYWORDS   HTG; HTGS_PHSSE2
SOURCE     Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
              clone:OSJNBA0082C09.
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1
AUTHORS    Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE      Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
            clone:OSJNBA0082C09
JOURNAL    Published Only in Database (2002)
REFERENCE  2 (bases 1 to 160562)
AUTHORS    Sasaki,T., Matsumoto,T. and Katayose,Y.

JOURNAL      Direct Submission
REFERENCE     Submitted (01-AUG-2002) Takuji Sasaki, National Institute of
              Agricultural Sciences, Rice Genome Research Program; Kannondai
              2-1-2, Tsukuba, Ibaraki 305-8602, Japan
              (E-mail:tsasakihias.affrc.go.jp, URL:http://rjp.dna.affrc.go.jp/,
              Tel:81-298-38-7441, Fax:81-298-38-7468)
              NOTE: It currently consists of 1 contigs. Gaps between the contigs
              are represented as runs of N. The order of the pieces is believed
              to be correct as given, however the sizes of the gaps between them
              are based on estimates that have provided by the submitter. This
              sequence will be replaced by the finished sequence as soon as it is
              available and the accession number will be preserved.
              * NOTE: This is a 'working draft' sequence.
              * This sequence will be replaced
              * by the finished sequence as soon as it is available and
              * the accession number will be preserved.
              Location/Qualifiers
              1. 160562
              /organism="Oryza sativa (japonica cultivar-group)"
              /cultivar="Nipponbare"
              /db_xref="taxon:39947"
              /clone="OSJNBA0082C09"
              /chromosome="2"
BASE COUNT   45817 a 33079 c 35854 g 45760 t 52 others
ORIGIN
Query Match      2.9%; Score 39; DB 2; Length 160562;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 847 TTCACGCGCGGTGCGACGACGCGCTGTGAGCAGTAC 885
|||||
DB 88249 TTCACGCGCGGTGCGACGACGCGCTGTGAGCAGTAC 88211
|||||

RESULT 3
AP003991/c 149417 bp DNA linear HTG 21-MAR-2002
LOCUS      Oryza sativa (japonica cultivar-group) chromosome 2 clone
DEFINITION CJ107_A12, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION  AP003991
VERSION     AP003991.1 GI:15076799
KEYWORDS   HTG; HTGS_PHSSE2.
SOURCE     Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
              clone:CJ107_A12.
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1
AUTHORS    Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE      Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
            clone:CJ107_A12
JOURNAL    Published Only in Database (2001)
REFERENCE  2 (bases 1 to 149417)
AUTHORS    Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE      Direct Submission
JOURNAL    Submitted (01-AUG-2001) Takuji Sasaki, National Institute of
              Agricultural Resources, Rice Genome Research Program; Kannondai
              2-1-2, Tsukuba, Ibaraki 305-8602, Japan
              (E-mail:tsasakihias.affrc.go.jp, URL:http://rjp.dna.affrc.go.jp/,
              Tel:81-298-38-7441, Fax:81-298-38-7468)
              NOTE: The nucleotide sequence of this BAC clone was generated by
              combining Monsanto and RGP-Japan sequencing data.
              COMBING: It currently consists of 1 contigs. Gaps between the contigs
              are represented as runs of N. The order of the pieces is believed
              to be correct as given, however the sizes of the gaps between them
              are based on estimates that have provided by the submitter. This
              sequence will be replaced by the finished sequence as soon as it is
              available and the accession number will be preserved.
              * NOTE: This is a 'working draft' sequence.
              * This sequence will be replaced
              * by the finished sequence as soon as it is available and
              * the accession number will be preserved.

```



```

FEATURES
  source
    * the accession number will be preserved.
    Location/Qualifiers
      1. 149417
      /organism="Oryza sativa (japonica cultivar-group)"
      /cultivar="Nipponbare"
      /db_xref="taxon:39947"
      /chromosome="2"
      /clone="OJ1077_A12"
BASE COUNT      41306 a 33053 c 33283 g 41528 t      247 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 4.7e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 596 CGCGGACATGACGCGCTGCGGCGCGCACAC 629
|||||
Db 42925 CGCGGACATGACGCGCTGCGGCGCGCACAC 42892
|||||

RESULT 4
LOCUS      BC005980      1496 bp      mRNA      linear      PRI 12-JUL-2001
DEFINITION Homo sapiens, ubiquitin-conjugating enzyme E2D 1 (homologous to
ACCESSION  BC005980
VERSION     BC005980.1 GI:13543662
KEYWORDS    MGC.
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1496)
REFERENCE   1
  AUTHORS   Strausberg, R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2001) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA
REMARK      NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT      Contact: MGC help desk
              Email: cgabs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Sequencing Group at the Stanford Human Genome
              Center, Stanford University School of Medicine, Stanford, CA 94305
              Web site: http://www.sbgc.stanford.edu
              Contact: (Dickson, Mark) mcdickpaxil.stanford.edu
              Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
              R. M.
  Clone distribution: MGC clone distribution information can be found
  through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
  Series: IRAL Plate: 21 Row: h Column: 13
  This clone was selected for full length sequencing because it
  passed the following selection criteria: Similarity but not
  identity to protein.
  Location/Qualifiers
    1. 1496
    /organism="Homo sapiens"
    /db_xref="LocusID:7321"
    /db_xref="taxon:9606"
    /clone="MGC:14673 IMAGE:4106710"
    /clone_type="Bone marrow, acute myelogenous leukemia"
    /clone_lib="NH_MGC_55"
    /lab_host="DH10B"
    /note="Vector: pDNR-LIB"
    /note="638
    195. .638
    /codon_start=1
    /product="ubiquitin-conjugating enzyme E2D 1 (homologous
    to yeast UBQ4/5)"
    /protein_id="AAH05980.1"

FEATURES
  source
    /db_xref="GI:13543663"
    /translation="MAKRIQKELSDLRDPPAHCSAGPYGDDLFHQATIMGPPSA
    YOGVEFLVHPPTDYPPRPPIAFETTKIIPNINSNGSLDLIRSQMSPALTVSKV
    LSTGSLCDPDPDDPDIAPDIAQIKSKKEKYNHNRMTQKVM"
BASE COUNT      461 a 311 c 261 g 463 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1322 CTTTTCGAAAAA
Db 1458 CTTTTCGAAAAA

RESULT 5
AP003544/c
LOCUS      AP003544      135792 bp      DNA      linear      HTG 21-MAR-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 6 clone P0599C12,
*** SEQUENCING IN PROGRESS ***. In ordered pieces.
ACCESSION  AP003544
VERSION     AP003544.1 GI:13810550
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
  clone:P0599C12
  Oryza sativa (japonica cultivar-group)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzaceae; Oryza.
  1
REFERENCE   1
  AUTHORS   Sasaki, T., Matsumoto, T. and Yamamoto, K.
  TITLE     Oryza sativa nipponbare(GAS) genomic DNA, chromosome 6, PAC
  clone:P0599C12
  JOURNAL   Published only in Database (2001)
  2 (bases 1 to 135792)
  Sasaki, T., Matsumoto, T. and Yamamoto, K.
  Direct Submission
  Submitted (25-APR-2001) Takuji Sasaki, National Institute of
  Agrobiological Resources, Rice Genome Research Program; Kannondai
  2-1-2, Tsukuba, Ibaraki 305-8602, Japan
  (E-mail:tsasaki@abrr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
  Tel:81-298-38-7441, Fax:81-298-38-7468)
  NOTE: It currently consists of 1 contigs. Gaps between the contigs
  are represented as runs of N. The order of the pieces is believed
  to be correct as given, however the sizes of the gaps between them
  are based on estimates that have provided by the submitter. This
  sequence will be replaced by the finished sequence as soon as it is
  available and the accession number will be preserved.
  * NOTE: This is a 'working draft' sequence.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
  Location/Qualifiers
    1. 135792
    /organism="Oryza sativa (japonica cultivar-group)"
    /cultivar="Nipponbare"
    /db_xref="taxon:39947"
    /chromosome="6"
    /clone="P0599C12"
BASE COUNT      37908 a 30304 c 30206 g 37124 t      250 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 847 TTCAACGGCGGCTGCGAGACGCGCTGTGAGC 879
|||||
Db 9334 TTCAACGGCGGCTGCGAGACGCGCTGTGAGC 9302
|||||

RESULT 6

```

AP004731 158456 bp DNA linear HTG-21-MAR-2002  
LOCUS AP004731  
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 6 clone  
OSJNBa0016D02, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.  
ACCESSION AP004731  
VERSION AP004731.1 GI:18656377  
KEYWORDS HTG-PHASE2  
SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,  
clone:OSJNBa0016D02.  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehretidae; Oryzae; Oryza.  
REFERENCE  
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC  
clone:OSJNBa0016D02  
JOURNAL Published Only in Database (2002)  
REFERENCE  
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
TITLE Direct Submission  
JOURNAL Submitted (13-FEB-2002) Takuji Sasaki, National Institute of  
Agricultural Sciences, Rice Genome Research Program; Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
NOTE: It currently consists of 1 contigs. Gaps between the contigs  
are represented as runs of 'N'. The order of the pieces is believed  
to be correct as given, however the sizes of the pieces between them  
are based on estimates that have provided by the submitter. This  
sequence will be replaced by the finished sequence as soon as it is  
available and the accession number will be preserved.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
FEATURES  
LOCATION/Qualifiers  
1..158456  
/organism="Oryza sativa (japonica cultivar-group)"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="6"  
/clone="OSJNBa0016D02"  
BASE COUNT 44183 a 35607 c 35786 g 42641 t 239 others  
ORIGIN  
Query Match 2.4%; Score 33; DB 2; Length 158456;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 847 TTCAACGGCGGGTCGACGACGCGCTGTGAGG 879  
|||||  
Db 92446 TTCAACGGCGGGTCGACGACGCGCTGTGAGG 92478  
RESULT 7  
AF384374 950 bp mRNA linear PLN 14-JUN-2001  
LOCUS AF384374  
DEFINITION Lycopersicon esculentum allene oxide cyclase (aoc) mRNA, complete  
cds  
ACCESSION AF384374  
VERSION AF384374.1 GI:14423350  
KEYWORDS Lycopersicon esculentum.  
SOURCE Lycopersicon esculentum  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE  
AUTHORS Weng,C.K. and To,K.Y.  
TITLE Molecular characterization of tomato allene oxide cyclase  
JOURNAL Unpublished

REFERENCE  
AUTHORS To,K.Y.  
TITLE Direct Submission  
JOURNAL Submitted (19-MAY-2001) Institute of Bioagricultural Sciences,  
Academia Sinica, Taipei 115, Taiwan  
Location/Qualifiers  
1..950  
/organism="Lycopersicon esculentum"  
/cultivar="CL5915"  
/db\_xref="taxon:4081"  
/clone="PR3-F21"  
/tissue\_type="root"  
1..950  
/gene="aoc"  
64..798  
/gene="aoc"  
/EC\_number="5.3.99.6"  
/function="jasmonic acid biosynthesis"  
/product="allene oxide cyclase"  
/protein\_id="AAK62358.1"  
/db\_xref="GI:14423351"  
/translation="MAVSSASALRTISSSSKLSAFOGKRIOSFRLPNPLISQNH  
KLTPTSTASRSFSCKSQSTSTSTVEQELSYEINERDGSFAYLRLOKTVNSL  
ADLVPSNKLVTDLKRIGITAGCIIKHEEKGRYEAIVSYFYGDGHAVOG  
GYLVEPEYLAIVGSGGIRPAGVGKRIQLLPFKLPFTYFKIGIPSLPDLCTAV  
PPEPTPEPPEAKCEGAAIKNTYN"  
BASE COUNT 294 a 220 c 174 g 262 t  
ORIGIN  
Query Match 2.4%; Score 32; DB 8; Length 950;  
Best Local Similarity 100.0%; Pred. No. 3e-06;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1323 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1354  
|||||  
Db 914 TTTTCGAAAAAAAAAAAAAAAAAAAAA 945  
RESULT 8  
AC019057 186896 bp DNA linear PRI 09-JAN-2002  
LOCUS AC019057  
DEFINITION Homo sapiens BAC clone RP11-121A1 from 2, complete sequence.  
ACCESSION AC019057  
VERSION AC019057.8 GI:15668119  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Sulston,J.E. and Waterston,R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 9847074  
PUBMED 9847074  
REFERENCE  
AUTHORS Tomlinson,C., Dixon,R., Kozlowski,A. and Boyer,E.  
TITLE The sequence of Homo sapiens BAC clone RP11-121A1  
JOURNAL Unpublished (2001)  
REFERENCE  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (30-DEC-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 186896)  
Waterston,R.H.  
Direct Submission  
JOURNAL Submitted (19-SEP-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 186896)  
REFERENCE

AUTHORS Waterston, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 COMMENT On Sep 19, 2001 this sequence version replaced g1:15144336.

Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@wustl.wustl.edu](mailto:sapiens@wustl.wustl.edu)  
 Summary Statistics  
 Center project name: H\_NH0121A01

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
 The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanese, J.V. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)  
 VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the left is RP11-814G20; the clone sequenced to the right is RP11-575E20, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-121A1; actual end is at base position 186896 of RP11-121A1.

Data from AC092170 and AC064870 was used to finish this clone, AC019057. Polymorphisms exist between RP11-121A1, RP11-814G20, and RP11-575E20.

FEATURES  
 source  
 1. 186896  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2"  
 /clone="RP11-121A1"  
 /clone\_11b="RPc1-11"  
 2296..2864  
 repeat\_region  
 /rpt\_family="L1"  
 2684..2713  
 repeat\_region  
 /rpt\_family="AT-rich"  
 3572..3621  
 repeat\_region  
 /rpt\_family="(TTG)n"  
 3631..3905  
 repeat\_region  
 /rpt\_family="Alu"  
 4339..4574  
 repeat\_region  
 /rpt\_family="MER1\_type"  
 5074..5123  
 repeat\_region  
 /rpt\_family="GA-rich"  
 5633..5654  
 repeat\_region

repeat\_region  
 /rpt\_family="AT-rich"  
 6075..6195  
 repeat\_region  
 /rpt\_family="MIR"  
 6727..6776  
 repeat\_region  
 /rpt\_family="MIR"  
 7332..7551  
 repeat\_region  
 /rpt\_family="MIR"  
 8361..8403  
 repeat\_region  
 /rpt\_family="L2"  
 9025..9325  
 repeat\_region  
 /rpt\_family="Alu"  
 9409..10233  
 misc\_feature  
 /note="match to EST BC674364 (NID:g13905760)"  
 9497..9580  
 repeat\_region  
 /rpt\_family="MIR"  
 9791..9886  
 repeat\_region  
 /rpt\_family="L2"  
 13134..13229  
 repeat\_region  
 /rpt\_family="CT-rich"  
 13974..14008  
 repeat\_region  
 /rpt\_family="(TAGA)n"  
 14050..14235  
 repeat\_region  
 /rpt\_family="(TA)n"  
 14236..14259  
 repeat\_region  
 /rpt\_family="(CA)n"  
 14309..14360  
 repeat\_region  
 /rpt\_family="(TA)n"  
 14450..14502  
 repeat\_region  
 /rpt\_family="(TA)n"  
 15191..15345  
 repeat\_region  
 /rpt\_family="L1"  
 15358..15647  
 repeat\_region  
 /rpt\_family="L1"  
 15474..15502  
 repeat\_region  
 /rpt\_family="AT-rich"  
 15648..15930  
 repeat\_region  
 /rpt\_family="Alu"  
 15931..16069  
 repeat\_region  
 /rpt\_family="L1"  
 16114..16527  
 repeat\_region  
 /rpt\_family="L1"  
 16529..16951  
 repeat\_region  
 /rpt\_family="L1"  
 16989..17375  
 repeat\_region  
 /rpt\_family="MAUR"  
 17584..17878  
 repeat\_region  
 /rpt\_family="L1"  
 18048..18073  
 repeat\_region  
 /rpt\_family="(A)n"  
 18355..18386  
 repeat\_region  
 /rpt\_family="(A)n"  
 19906..21166  
 repeat\_region  
 /rpt\_family="L1"  
 20055..20080  
 repeat\_region  
 /rpt\_family="(CA)n"  
 20413..20438  
 repeat\_region  
 /rpt\_family="(GAGA)n"  
 21109..21137  
 repeat\_region  
 /rpt\_family="AT-rich"  
 21161..21182  
 repeat\_region  
 /rpt\_family="AT-rich"  
 21952..22122  
 repeat\_region  
 /rpt\_family="MER1\_type"  
 21975..22319  
 misc\_feature  
 /note="match to EST AV748686 (NID:g10906534)"  
 22625..23194  
 repeat\_region  
 /rpt\_family="ERV1"  
 23634..23759  
 repeat\_region  
 /rpt\_family="L2"  
 24006..24310  
 repeat\_region  
 /rpt\_family="L2"  
 25275..25329  
 repeat\_region  
 /rpt\_family="GA-rich"

```

repeat_region 25682..25716
                /rpl_family="(TTTTTC)n"
repeat_region 26161..26212
                /rpl_family="(TTTAA)n"
repeat_region 26210..26493
                /rpl_family="Alu"
repeat_region 28293..29275
                /rpl_family="L1"
repeat_region 29272..29416
                /rpl_family="L1"
repeat_region 29449..29549
                /rpl_family="L2"
repeat_region 29550..29843
                /rpl_family="Alu"
repeat_region 29844..30447
                /rpl_family="L2"
repeat_region 30938..31268
                /rpl_family="Mariner"
repeat_region 32065..33222
                /rpl_family="MIR"
repeat_region 32669..32706
                /rpl_family="AT-rich"

```

Query Match 2.4%; Score 32; DB 9; Length 186896;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-06;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1323 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
 DB 18040 TTTTCGAAAAAAAAAAAAAAAAAAAA 18071

RESULT 9 AK026389  
 LOCUS Homo sapiens cDNA: FLJ27236 fls, clone HUV00250, highly similar to  
 AF131821 Homo sapiens clone 24877 mRNA sequence.  
 ACCESSION AK026389  
 VERSION AK026389.1 GI:10439239  
 KEYWORDS Oligo capping; fls (full insert sequence).  
 SOURCE Homo sapiens primary endothelial cells of human umbilical vein cDNA  
 to mRNA, clone lib: HUV clone: HUV00250.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,  
 Matsunaga,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,  
 Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hiroo,M.,  
 Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,  
 Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 1253)  
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,  
 Shibahara,T., Tanaka,T. and Nakamura,Y.

TIME Direct Submission  
 JOURNAL Submitted (29-NOV-2000) Sumio Sugano, Institute of Medical Science,  
 University of Tokyo, Laboratory of Genome Structure Analysis, Human  
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,  
 Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
 Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan; cDNA full insert  
 sequencing: Research Association for Biotechnology; cDNA library  
 construction: 3 - 6 3 - end one pass sequencing; Departent of  
 Virology and Human Genome Center, Institute of Medical Science,  
 University of Tokyo (partly supported by Science and Technology  
 Agency).

FEATURES  
 source 1..1253  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

```

/cclone="HUV00250"
/cell_type="Primary endothelial cells of human umbilical
vein"
/clone_lib="HUV"
/clone="cloning vector pMF185FL3"
misc-feature 1..1253
              /note="highly similar to AF131821 Homo sapiens clone 24877
              mRNA sequence"
BASE COUNT 347 a 346 c 300 g 260 t
ORIGIN

```

Query Match 2.3%; Score 31; DB 9; Length 1253;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
 DB 1196 TTTTCGAAAAAAAAAAAAAAAAAAAA 1226

RESULT 10 HSM804954  
 LOCUS HSM804954 1495 bp mRNA linear PRI 10-JUL-2002  
 DEFINITION Homo sapiens mRNA; cDNA DKFZp686M0979 (from clone DKFZp686M0979).  
 ACCESSION AL833641  
 VERSION AL833641.1 GI:21734288  
 KEYWORDS

SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Well,B., Amid,C. and  
 Wiemann,S.  
 TITLE Direct Submission  
 JOBNRL Submitted (09-JUL-2002) 1, D-85764 Neuberberg, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
 consortium of the German Genome Project.  
 This clone (DKFZp686M0979) is available at the R2PD in Berlin.  
 Please contact the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@r2pd.de Further  
 information about the clone and the sequencing project is available  
 at http://mips.gsf.de/proj/cDNA/.

FEATURES  
 source 1..1495  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp686M0979"  
 /tissue\_type="cDNA-collection"  
 /clone\_lib="686 (synonym: hlc3). Vector pSport1\_Sfl; host  
 DH10B; sites SfiI + SfiIb"  
 /dev\_stage="adult"

polyA\_signal 1422..1427  
 polyA\_site 1444  
 BASE COUNT 388 a 414 c 390 g 303 t  
 ORIGIN

Query Match 2.3%; Score 31; DB 9; Length 1495;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1323 TTTTCGAAAAAAAAAAAAAAAAAAAA 1353  
 DB 1437 TTTTCGAAAAAAAAAAAAAAAAAAAA 1467

RESULT 11 AX055708  
 LOCUS AX055708 1844 bp DNA linear PAT 13-JAN-2001  
 DEFINITION Sequence 23 from Patent WO0073348.  
 ACCESSION AX055708

```

VERSION      AX055708.1  GI:12228837
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
REFERENCE      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS      Baker,K.P., Goddard,A., Gurney,A.L., Hebert,C., Henzel,W.,
              Khabaroff,R.C., Shelton,D.L., Smith,V., Watanabe,C.K. and Wood,W.I.
              Methods and compositions for inhibiting neoplastic cell growth
              Patent: WO 0073348-A 23 07-DEC-2000;
TITLE
JOURNAL
FEATURES
  source
    1. .1844
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT    398 a 564 c 590 g 292 t
ORIGIN

Query Match
  Best Local Similarity 100.0%; Pred. No. 1.2e-05;
  Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

RESULT 12
LOCUS      AX089940      1844 bp      DNA      linear      PAT 21-MAR-2001
DEFINITION Sequence 1 from Patent WO0116319.
ACCESSION  AX089940
VERSION     AX089940.1  GI:13443981
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
              1 (bases 1 to 1844)
AUTHORS      Goddard,A., Godowski,P.J., Gurney,A.L., Hillan,K.J., Tumas,D.,
              Watanabe,C.K. and Wood,W.I.
              Compositions and methods for the treatment of immune related
              diseases
              Patent: WO 0116319-A 1 08-MAR-2001;
TITLE
JOURNAL
FEATURES
  source
    1. .1844
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT    398 a 564 c 590 g 292 t
ORIGIN

Query Match
  Best Local Similarity 100.0%; Pred. No. 1.2e-05;
  Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

RESULT 13
LOCUS      AX454590      1844 bp      DNA      linear      PAT 06-JUL-2002
DEFINITION Sequence 175 from Patent WO0208284.
ACCESSION  AX454590
VERSION     AX454590.1  GI:21713919
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```

```

REFERENCE      1 Baker,K.P., Ferrara,N., Gerber,H., Gerritsen,M.E., Goddard,A.,
              Godowski,P.J., Gurney,A.L., Hillan,K.J., Marsters,S.A., Pan,J.,
              Paoni,N.F., Stephan,J.P., Watanabe,C.K., Williams,P.M., Wood,W.I.
              and Ye,W.
              Compositions and methods for the diagnosis and treatment of
              disorders involving angiogenesis
              Patent: WO 0208284-A 175 31-JAN-2002;
TITLE
JOURNAL
FEATURES
  source
    1. .1844
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT    398 a 564 c 590 g 292 t
ORIGIN

Query Match
  Best Local Similarity 100.0%; Pred. No. 1.2e-05;
  Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

RESULT 14
LOCUS      AX464342      1844 bp      DNA      linear      PAT 16-JUL-2002
DEFINITION Sequence 475 from Patent WO0140466.
ACCESSION  AX464342
VERSION     AX464342.1  GI:21899187
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
              1
AUTHORS      Baker,K.P., Beresini,M., DeForge,L., Desnoyers,L., Filvaroff,E.,
              Go,W.O., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
              Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
              Wood,W.L. and Zhang,Z.
              Secreted and transmembrane polypeptides and nucleic acids encoding
              same
              Patent: WO 0140466-A 475 07-JUN-2001;
TITLE
JOURNAL
FEATURES
  source
    1. .1844
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT    398 a 564 c 590 g 292 t
ORIGIN

Query Match
  Best Local Similarity 100.0%; Pred. No. 1.2e-05;
  Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

RESULT 15
LOCUS      AX491068      1844 bp      DNA      linear      PAT 16-AUG-2002
DEFINITION Sequence 175 from Patent WO0200690.
ACCESSION  AX491068
VERSION     AX491068.1  GI:22323871

```

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1  
Baker, K.P., Ferrara, N., Gerber, H., Geritsen, M.E., Goddard, A.,  
Godowski, P.J., Gurney, A.L., Hillan, K.J., Marsters, S.A., Pan, J.,  
Pacini, N.F., Stephan, J.P., Watanabe, C.K., Williams, P.M., Wood, W.I.  
and Ye, W.  
Compositions and methods for the diagnosis and treatment of  
disorders involving angiogenesis  
Patent: WO 0200690-A 175 03-JAN-2002;  
Genentech, Inc. (US)

Location/Qualifiers  
1. 1844  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

398 a 564 c 590 g 292 t

Query Match 2.3%; Score 31; DB 6; Length 1844;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
|||||  
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817  
|||||

Search completed: May 3, 2003, 11:14:25  
Job time : 3317 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 09:24:46 ; Search time 243 Seconds  
(Without alignments)  
12548.172 Million cell updates/sec

Title: US-10-047-825-3

Perfect score: 1354  
Sequence: 1 aatcgcgcagcagcttaagc.....aaaaaaaaaaaaaaaaaa 1354

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 segs, 112599159 residues

Word size: 16

Total number of hits satisfying chosen parameters: 101640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N\_Geneseq\_101002.\*

```
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	32	2.4	295	23	ABV05070 Human prostate exp
C 2	32	2.4	369	23	ABV35332 Human prostate exp
C 3	32	2.4	428	23	ABV14239 Human prostate exp
4	32	2.4	1527	21	AAC98823 Human pancreatic c
5	31	2.3	1722	21	AAZ98108 Human secreted pro
6	31	2.3	1722	22	AAAD11721 Human secreted pro
7	31	2.3	1722	24	ABK69618 Human secreted pro
8	31	2.3	1843	21	AAA37046 Human PRO1265 (UNO
9	31	2.3	1844	20	AAK87257 cDNA clone encodin

10	31	2.3	1844	21	AAA46906 cDNA encoding nove
11	31	2.3	1844	22	AAK21481 Human cDNA sequenc
12	31	2.3	1844	22	AAAD02920 Human PRO1265 cDNA
13	31	2.3	1844	22	AAAF54260 DNA encoding prote
14	31	2.3	1844	22	AAC91564 Human PRO1265 cDNA
15	31	2.3	1844	22	ABL95648 Human angiogenesis
16	31	2.3	1844	24	ABR88159 Human PRO1265 cDNA
17	31	2.3	1867	21	AAZ98087 Human secreted pro
18	31	2.3	1867	22	AAAD11700 Human secreted pro
19	31	2.3	1867	22	ABK69796 Human secreted pro
20	31	2.3	3026	24	AAI66970 Human ion channel
21	31	2.3	5964	24	ABN80171 Human chemically m
22	30	2.2	121	19	AAV54078 Nucleotide sequenc
23	30	2.2	285	23	ABV60247 Human prostate exp
24	30	2.2	2086	22	AAAD05733 Male ferulate-5-h
25	30	2.2	2900	21	AAAF9688 Pinus radiata cell
26	29	2.1	446	22	AAAL10267 Human breast cance
27	29	2.1	1343	20	AAAX9858 Human secreted pro
28	29	2.1	1343	22	AAAS9282 Human cDNA encodin
29	29	2.1	1343	24	ABR90851 Human polynucleoti
30	29	2.1	8294	24	ABL34431 Human immune syste
31	29	2.1	13376	24	ABL32582 Human immune syste
32	28	2.1	196	24	ABO56551 Human colon cancer
33	28	2.1	299	20	AAZ14537 Human gene express
34	28	2.1	300	20	AAZ14536 Stomach cancer rel
35	28	2.1	302	24	ABL64171 Human prostate exp
36	28	2.1	314	23	ABV47595 Human colon cancer
37	28	2.1	442	24	ABQ56852 Human secreted pro
38	28	2.1	608	21	AAAC68087 Human ovarian anti
39	28	2.1	698	24	ABQ54353 Human secreted pro
40	28	2.1	705	24	ABL87293 Human secreted pro
41	28	2.1	952	22	AAH31422 Human ORFX ORF2726
42	28	2.1	1184	21	AACT7171 Human secreted pro
43	28	2.1	1119	24	ABQ54352 Human secreted pro
44	28	2.1	1508	20	AAK20416 Human P40 encoding
45	28	2.1	2024	24	ABN85573

#### ALIGNMENTS

RESULT 1	ABV05070/c	ABV05070 standard; cDNA: 295 BP.
XX	ABV05070;	
AC	13-SEP-2002 (first entry)	
XX	Human prostate expression marker cDNA 5061.	
DE	Human; prostate cancer; cytostatic; carcinogen; pharmacodynaic marker;	
KW	pharmacogenomic marker; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200160860-A2.	
XX		
PD	23-AUG-2001.	
XX		
PF	20-FEB-2001; 2001WO-US05171.	
XX		
PR	17-FEB-2000; 2000US-183319P.	
PR	16-MAR-2000; 2000US-189862P.	
PR	25-MAY-2000; 2000US-207454P.	
PR	09-JUN-2000; 2000US-211314P.	
PR	18-JUL-2000; 2000US-219007P.	
PR	13-DEC-2000; 2000US-255281P.	
XX		
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
PI	Schlegel R, Endege WO, Monahan JE;	
XX		

DR WPI: 2001-662795/76.  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer  
 XX  
 PS Claim 1; Page 861; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 CC  
 XX Sequence 295 BP; 72 A; 58 C; 58 G; 92 T; 15 other;  
 SQ  
 Query Match 2.4%; Score 32; DB 23; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 0.0056;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1323 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
 DB 40 TTTTCGAAAAAAAAAAAAAAAAAAAA 9  
 RESULT 2  
 ABV35332/c  
 ID ABV35332 standard; cDNA; 369 BP.  
 XX  
 AC ABV35332;  
 XX  
 DT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 35323.  
 XX  
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KM pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US05171.  
 XX  
 PR 17-FEB-2000; 2000US-183319P.  
 PR 16-MAR-2000; 2000US-189862P.  
 PR 25-MAY-2000; 2000US-207454P.  
 PR 09-JUN-2000; 2000US-211314P.  
 PR 18-JUL-2000; 2000US-219007P.  
 PR 13-DEC-2000; 2000US-255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PI Schlegel R, Endege WO, Monahan JE;  
 PT  
 DR WPI: 2001-662795/76.  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer -  
 XX

PS Claim 1; Page 7364; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 CC  
 XX Sequence 369 BP; 88 A; 85 C; 88 G; 108 T; 0 other;  
 SQ  
 Query Match 2.4%; Score 32; DB 23; Length 369;  
 Best Local Similarity 100.0%; Pred. No. 0.0056;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1323 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
 DB 72 TTTTCGAAAAAAAAAAAAAAAAAAAA 41  
 RESULT 3  
 ABV14239/c  
 ID ABV14239 standard; cDNA; 428 BP.  
 XX  
 AC ABV14239;  
 XX  
 DT 13-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 14230.  
 XX  
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KM pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US05171.  
 XX  
 PR 17-FEB-2000; 2000US-183319P.  
 PR 16-MAR-2000; 2000US-189862P.  
 PR 25-MAY-2000; 2000US-207454P.  
 PR 09-JUN-2000; 2000US-211314P.  
 PR 18-JUL-2000; 2000US-219007P.  
 PR 13-DEC-2000; 2000US-255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PI Schlegel R, Endege WO, Monahan JE;  
 PT  
 DR WPI: 2001-662795/76.  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer -  
 XX  
 PS Claim 1; Page 2374; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 CC  
 XX Sequence 369 BP; 88 A; 85 C; 88 G; 108 T; 0 other;  
 SQ  
 Query Match 2.4%; Score 32; DB 23; Length 369;  
 Best Local Similarity 100.0%; Pred. No. 0.0056;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1323 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
 DB 72 TTTTCGAAAAAAAAAAAAAAAAAAAA 41  
 RESULT 3  
 ABV14239/c  
 ID ABV14239 standard; cDNA; 428 BP.  
 XX  
 AC ABV14239;  
 XX  
 DT 13-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 14230.  
 XX  
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KM pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US05171.  
 XX  
 PR 17-FEB-2000; 2000US-183319P.  
 PR 16-MAR-2000; 2000US-189862P.  
 PR 25-MAY-2000; 2000US-207454P.  
 PR 09-JUN-2000; 2000US-211314P.  
 PR 18-JUL-2000; 2000US-219007P.  
 PR 13-DEC-2000; 2000US-255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PI Schlegel R, Endege WO, Monahan JE;  
 PT  
 DR WPI: 2001-662795/76.  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer -  
 XX



(b) monitoring the progression of prostate cancer in a patient;  
(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;  
(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;  
(e) selecting a composition for inhibiting prostate cancer in a patient;  
(f) assessing the prostate cell carcinogenic potential of a compound;  
(g) determining whether prostate cancer has metastasized in a patient;  
(h) assessing the aggressiveness or indolence of prostate cancer in a patient;  
(i) is also useful as a pharmacodynamic or pharmacogenomic marker.

Query Match 2.4%; Score 32; DB 23; Length 428;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1323 TTTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
Db 47 TTTTTCGAAAAAAAAAAAAAAAAAAAA 16

## RESULT 4

AAC98823  
ID AAC98823 standard; cDNA; 1527 BP.

AAC98823;

09-MAR-2001 (first entry)

Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:51.

Human; pancreas; pancreatic cancer; pancreatic cancer antigen;  
detection; diagnosis; identification; cytosolic; neuroprotective;  
neurotropic; immunomodulatory; relaxant; contraceptive; gynaecological;  
antiinflammatory; cardiac; gene therapy; chromosome mapping;  
linkage analysis; tissue identification; tissue typing; forensic;  
neutral; immune system; muscular; reproductive; gastrointestinal;  
pulmonary; cardiovascular; renal; proliferative; ss.

Homo sapiens.

WO200055320-A1.

21-SEP-2000.

08-MAR-2000; 2000WO-US05989.

12-MAR-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI: 2000-579444/54.

P-PSDB; AAB54058.

New nucleic acid that is a pancreatic cancer antigen for preventing,  
treating, or ameliorating a medical condition, particular pancreatic  
cancer, or for use in assays for diagnosing a pathological condition -  
Claim 1; Page 531-532; 1379pp; English.

AAC98773 to AAC99231 encode the human pancreatic cancer associated  
proteins, called pancreatic cancer antigens, given in AAB54008 to  
AAB54466. The human pancreatic cancer antigens have cytosolic,  
neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive,  
gynaecological, cardiac and antiinflammatory activities, and can be used  
in gene therapy. The polynucleotide and proteins can be used for  
preventing, treating, or ameliorating a medical condition or in assays  
for diagnosing a pathological condition or a susceptibility to one in a  
subject. Binding partners to the proteins and the activity of the

proteins can be identified. The pancreatic cancer antigens can be used to  
detect, treat or prevent pancreatic disorders, especially cancer.  
Agonists and antagonists to the antigens can be screened for. The  
pancreatic cancer antigen polynucleotides can be used to design nucleic  
acid hybridisation probes that can be used in chromosome mapping, linkage  
analysis, tissue identification and/or typing and a variety of forensic  
and diagnostic methods. The proteins can be used to generate antibodies  
which are used to purify, detect and target the polypeptides, including  
both in vivo and in vitro diagnostic and therapeutic methods. The  
proteins can be used to treat or prevent neutral, immune system, muscular,  
reproductive, gastrointestinal, pulmonary, cardiovascular, renal or  
proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent  
sequences used in the exemplification of the present invention.

Query Match 2.4%; Score 32; DB 21; Length 1527;  
Best Local Similarity 100.0%; Pred. No. 0.0043;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1323 TTTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
Db 1487 TTTTTCGAAAAAAAAAAAAAAAAAAAA 1518

## RESULT 5

AA298108  
ID AA298108 standard; cDNA; 1722 BP.

AA298108;

09-MAY-2000 (first entry)

Human secreted protein encoding nucleotide sequence SEQ ID NO:102.

Human; secreted protein; diagnosis; cytosolic; immunosuppressive;  
antiinflammatory; neurotropic; neuroprotective; antiallergic; cancer;  
tumour; neurodegenerative disorder; developmental abnormality; allergy;  
foetal deficiency; blood disorder; immune system disorder; arthritis;  
autoimmune disease; hepatic disease; renal disease; inflammation;  
Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;  
infection; AIDS; spinal cord injury; transplant rejection; diabetes;  
asthma; sepsis; acne; psoriasis; cardiovascular disorder;  
reproductive disorder; gastrointestinal disorder; respiratory disorder;  
metabolic disorder; food additive; preservative; ss.

Homo sapiens.

WO200004140-A1.

27-JAN-2000.

14-JUL-1999; 99WO-US15849.

15-JUL-1998; 98US-0092921.

PR 15-JUL-1998; 98US-0092922.

15-JUL-1998; 98US-0092956.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;  
Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;  
Mucenski M, Endress GA, Soppet DR;

WPI: 2000-161128/14.

P-PSDB; AAY87155.

New isolated human genes, useful for diagnosis and treatment of, e.g.  
cancers, neurological or blood disorders -

Claim 1; Page 370-371; 494pp; English.

The polynucleotide sequences given in AA298017 to AA298108 encode the

human secreted proteins given in AAY87064 to AAY87223. Human secreted protein can have activities based on the tissues and cells the genes are expressed in. Examples of activities include: cytostatic; immunosuppressive; antiinflammatory; neurotrophic; neuroprotective; and antiallergic. The polynucleotides and their corresponding secreted polypeptides are useful for preventing, treating or ameliorating medical conditions, e.g. by protein or gene therapy. Also pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new polynucleotides. Human secreted proteins and their polynucleotides can be used for developing products for the diagnosis or treatment of cancer, tumours, neurodegenerative disorders, developmental abnormalities and foetal deficiencies, blood disorders, diseases of the immune system, autoimmune diseases, hepatic and renal disease, inflammation, allergies, Alzheimer's disease, behavioural disorders, schizophrenia, osteoporosis, arthritis, infections, AIDS, spinal cord injuries, transplant rejection, diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders, reproductive disorders, gastrointestinal disorders, respiratory disorders and metabolic disorders. The proteins or polynucleotides can also be used as food additives or preservatives. The proteins are also useful for identifying their binding partners. AA298008 to AA298016 and AAY87063 are sequence used in the exemplification of the present invention.

Sequence 1722 BP: 376 A; 534 C; 538 G; 269 T; 5 other:  
Query Match 2.3%; Score 31; DB 22; Length 1722;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1324 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1354  
1655 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1685

RESULT 6  
AAD11721  
AAD11721 standard; cDNA: 1722 BP.

AAD11721;

24-SEP-2001 (first entry)

Human secreted protein-encoding gene 71 cDNA clone HDP0256, SEQID NO:102.

Human; secreted protein; proliferative disorder; cancer; tumour; asthma; foetal abnormality; developmental abnormality; haematoid arthritis; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; Parkinson's disease; cognitive disorder; schizophrenia; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; inflammation; neurological disorder; Alzheimer's disease; food additive; angiotensin disorder; kidney disorder; gastrointestinal disorder; allergy; pregnancy-related disorder; endocrine disorder; infection; wound healing; cell culture; chemotaxis; vulnery; binding partner identification; gene therapy; ss.

Homo sapiens.

Location/Qualifiers  
key 59..1018  
CDS

Product= "Human secreted protein precursor"  
Transl\_except= (pos:260..262, aa:Xaa)  
Transl\_except= (pos:401..403, aa:Xaa)  
Transl\_except= (pos:695..697, aa:Xaa)  
59..121  
tag= b  
tag= 122..1015  
tag= C  
Product= "Mature human secreted protein"

MO200151504-A1.

19-JUL-2001.  
12-JAN-2001: 2001WO-US000911.  
13-JAN-2000: 2000US-0482273.  
(HUMA-) HUMAN GENOME SCI INC.  
Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y, Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR, Endress GA, Muscenski M, Ebner R, WPI, 2001-425865/45.  
P-PSDB: AAE06132.  
Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition  
Claim 1: Page 737-738; 864pp; Eng11sh.

AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted protein genes, and AAE06041-AAE06132 represent the proteins they encode. AAE06133-AAE06205 represent human secreted protein fragments. The secreted proteins and their genes are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 71 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiotensin disorders, kidney disorders, endocrine disorders, and infections. The pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays (e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA)). The present sequence represents a human secreted protein-encoding cDNA of the invention.

Sequence 1722 BP: 376 A; 534 C; 538 G; 269 T; 5 other:  
Query Match 2.3%; Score 31; DB 22; Length 1722;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1324 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1354  
1655 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1685

RESULT 7  
ABK69818  
ID ABK69818 standard; cDNA: 1722 BP.

ABK69818;

15-JUL-2002 (first entry)

Human secreted protein gene 71 #2.

Human; ss; gene; secreted protein; gene therapy; immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;





PT Antibody against proteins expressed in neoplastic cells, useful for  
PT tumor diagnosis and treatment

PS Example 1; Fig 7; 162pp; English.

XX This is the nucleotide sequence of cDNA clone DNA60764 (ATCC 203452)  
CC coding for human PRO1265 (UN0636) (see AAY06480). The clone was  
CC isolated from an adenoid tissue library. Amplification of DNA60764  
CC on chromosome 19 occurs in various lung and colon tumours and cell  
CC lines, suggesting a significant role in tumour formation and growth.  
CC Antagonists (e.g. antibodies) directed tot PRO1265 are expected to  
CC find use in cancer therapy. The invention identifies 14 genes (see  
CC A487254-67) that are amplified in the genome of tumour cells. Such  
CC amplification is expected to be associated with overexpression of  
CC the gene product and to contribute to tumorigenesis. The encoded  
CC proteins (see AAY06477-90) may be useful targets for the diagnosis  
CC and/or treatment (including prevention) of certain cancers, and may  
CC act as predictors of the prognosis of tumour treatment.

XX Sequence 1844 BP; 398 A; 564 C; 590 G; 292 T; 0 other;

Query Match 2.3%; Score 31; DB 20; Length 1844;

Best Local Similarity 100.0%; Pred. No. 0.01; Mismatches 0; Gaps 0;

Matches 31; Conservative 0; Indels 0; Gaps 0;  
OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1354  
DB 1787 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1817

RESULT 10

AAA46906  
ID AAA46906 standard; cDNA; 1844 BP.

AC AAA46906;

DT 03-OCT-2000 (first entry)

XX cDNA encoding novel polypeptide PRO1265.

DE PRO201; PRO292; PRO327; PRO1265; PRO344; PRO343; PRO347; PRO357;  
KW PRO15; PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell;  
KW tumorigenesis; cancer; neoplastic cell growth; cell proliferation; ss.

OS Homo sapiens.

PH Key Location/Qualifiers  
FT CDS 79..1782  
FT /\*tag= a

XX WO200037640-A2.

XX 29-JUN-2000.

XX 16-DEC-1999; 99WO-US30095.

XX 22-DEC-1998; 98US-0113296.

XX 08-MAR-1999; 99WO-US05028.

XX 02-JUN-1999; 99WO-US12252.

XX 01-SEP-1999; 99WO-US20111.

XX 15-SEP-1999; 99WO-US21090.

XX 30-NOV-1999; 99WO-US28313.

XX 30-NOV-1999; 99WO-US28409.

XX 01-DEC-1999; 99WO-US28301.

XX 02-DEC-1999; 99WO-US28565.

XX (GENTH) GENTECH INC.

XX Botstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA;

PI Wood W;

XX WPI; 2000-452188/39.

DR P-PSDB; AAY93687.

XX New anti-polypeptide antibody useful in the treatment and diagnosis of  
PT neoplastic cell growth and proliferation -

PS Claim 50; Fig 7; 220pp; English.

XX The present sequence encodes a novel human polypeptide. The  
CC specification describes novel polypeptides designated PRO201, PRO292,  
CC PRO327, PRO1265, PRO344, PRO347, PRO357, PRO1017,  
CC PRO1112, PRO509, PRO853 and PRO882. These genes are amplified in  
CC the genome of tumour cells. The polypeptides are believed to contribute  
CC to tumourigenesis. The polypeptides are useful target for the  
CC identification of certain cancers, and may act as predictors of the  
CC prognosis of tumour treatment. Antibodies against these polypeptides  
CC are useful in the treatment and diagnosis of neoplastic cell growth  
CC and proliferation in mammals.

XX Sequence 1844 BP; 398 A; 564 C; 590 G; 292 T; 0 other;

Query Match 2.3%; Score 31; DB 21; Length 1844;

Best Local Similarity 100.0%; Pred. No. 0.01; Mismatches 0; Gaps 0;

Matches 31; Conservative 0; Indels 0; Gaps 0;  
OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1354  
DB 1787 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1817

RESULT 11

AAS21481  
ID AAS21481 standard; cDNA; 1844 BP.

AC AAS21481;

DT 24-OCT-2001 (first entry)

XX Human cDNA sequence encoding for PRO1265 polypeptide.

DE Human secretory and transmembrane; PRO; mammalian; cancer; lung;

KW breast; prostate; cervical; tumour necrosis factor- $\alpha$ 1pha; TNF- $\alpha$ 1pha;

KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;

KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.

OS Homo sapiens.

PH WO200140466-A2.

XX 07-JUN-2001.

XX 01-DEC-2000; 2000WO-US32678.

XX 01-DEC-1999; 99WO-US28301.

XX 01-DEC-1999; 99WO-US28634.

XX 02-DEC-1999; 99WO-US28565.

XX 02-DEC-1999; 99WO-US28564.

XX 09-DEC-1999; 99US-0170262.

XX 16-DEC-1999; 99WO-US30095.

XX 20-DEC-1999; 99WO-US30911.

XX 30-DEC-1999; 99WO-US31243.

XX 06-JAN-2000; 2000WO-US00377.

XX 11-FEB-2000; 2000WO-US03565.

XX 18-FEB-2000; 2000WO-US04341.

XX 22-FEB-2000; 2000WO-US04342.

XX 24-FEB-2000; 2000WO-US04914.

XX 24-FEB-2000; 2000WO-US05004.

XX 01-MAR-2000; 2000WO-US05601.

XX 20-MAR-2000; 2000WO-US07377.

XX 21-MAR-2000; 2000WO-US07532.

XX 30-MAR-2000; 2000WO-US08439.



Sequence 1844 BP; 398 A; 564 C; 590 G; 292 T; 0 other;

Query Match 2.3%; Score 31; DB 22; Length 1844;

Best Local Similarity 100.0%; Pred. No. 0.01; Mismatches 0; Indels 0; Gaps 0;

DB 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

# RESULT 13

AAFS4260  
ID AAF54260 standard; DNA; 1844 BP.

AAFS4260;

02-APR-2001 (first entry)

DNA encoding protein of the invention #25.

Secreted; transmembrane; gene therapy; ss.

Unidentified.

WO200078961-A1.

28-DEC-2000.

18-FEB-2000; 2000WO-US04342.

23-JUN-1999; 99US-0141037.

20-JUL-1999; 99US-0144758.

26-JUL-1999; 99US-0145698.

01-SEP-1999; 99WO-US20111.

29-OCT-1999; 99WO-US28313.

30-NOV-1999; 99WO-US28551.

02-DEC-1999; 99WO-US30095.

16-DEC-1999; 99WO-US30095.

05-JAN-2000; 2000WO-US00219.

06-JAN-2000; 2000WO-US00376.

(GETH ) GENENTECH INC.

Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S, Geo W, Goddard A, Godowski FJ, Grimaldi CJ, Gurney AL, Hillan KJ, Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;

WPI; 2001-071395/08.

Secreted and transmembrane proteins and nucleic acids designated PRO, useful as hybridization probes, in chromosome and gene mapping and gene therapy -

Claim 2; Fig 49; 787bp; English.

The present invention relates to secreted and transmembrane proteins. These proteins and the DNA encoding them may be used as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. They may also be used to generate either transgenic animals or knockout animals which are in turn useful for development and screening of therapeutically useful reagents.

The nucleic acids may also be used in gene therapy.

Sequence 1844 BP; 398 A; 564 C; 590 G; 292 T; 0 other;

Query Match 2.3%; Score 31; DB 22; Length 1844;

Best Local Similarity 100.0%; Pred. No. 0.01; Mismatches 0; Indels 0; Gaps 0;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

DB 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

# RESULT 14

AAAC91564  
ID AAC91564 standard; cDNA; 1844 BP.

AAAC91564;

21-MAR-2001 (first entry)

Human PRO1265 cDNA.

Human; PRO; cytosolic; neurotrophic; neuroprotective; respiratory general; antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant; PRO agonist; cancer; inflammatory disorder; immunological disorder; ss.

Homo sapiens.

WO200073348-A2.

07-DEC-2000.

30-MAY-2000; 2000WO-US14941.

02-JUN-1999; 99WO-US12252.

22-JUN-1999; 99US-0140650.

23-JUN-1999; 99US-0141037.

20-JUL-1999; 99US-0144758.

01-SEP-1999; 99WO-US20111.

08-SEP-1999; 99WO-US20594.

29-OCT-1999; 99US-0162806.

30-NOV-1999; 99WO-US28313.

01-DEC-1999; 99WO-US28551.

02-DEC-1999; 99WO-US30095.

16-DEC-1999; 99WO-US30095.

20-DEC-1999; 99WO-US30095.

06-JAN-2000; 2000WO-US00376.

11-FEB-2000; 2000WO-US03565.

18-FEB-2000; 2000WO-US04341.

02-MAR-2000; 2000WO-US05841.

03-MAR-2000; 2000US-0187202.

10-MAR-2000; 2000WO-US06319.

15-MAR-2000; 2000WO-US06884.

30-MAR-2000; 2000WO-US08439.

17-MAY-2000; 2000WO-US13705.

(GETH ) GENENTECH INC.

Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC, Shelton DL, Smith V, Watanabe CK, Wood WI;

P-PSDB; AAB50962.

WPI; 2001-016509/02.

Twenty eight nucleic acids encoding PRO polypeptides which are useful for treating various tumors, e.g. breast cancer, and other inflammatory, angiogenic and immunological disorders -

Claim 20; Fig 23; 188bp; English.

The present sequence is one of twenty eight nucleic acids encoding PRO polypeptides. The PRO polypeptides and their agonists, including antibodies, peptides, and small molecule agonists, may be used to treat various tumors, e.g., cancers such as breast cancer, ovarian cancer, lung cancer, renal cancer, colorectal cancer, uterine cancer, prostate cancer, leukemia, bladder cancer, central nervous system cancer, melanoma or leukadema. They are also useful for treating other disorders such as neuronal, glial, astrocytic, hypothalamic and other glandular, macrophagal, epithelial, stromal and blastocoeleic disorders, and inflammatory, angiogenic and immunological disorders.





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 10:12:17 ; Search time 50 Seconds  
(without alignments)  
8304.807 Million cell updates/sec

Title: US-10-047-825-3

Perfect score: 1354  
Sequence: 1 aattcgacgagcttaagc.....aaaaaaaaaaaaaaaa 1354

Scoring table: OLIGO.MUC  
Gapop 60.0 , Gapext 60.0

Searched: 41362 seqs, 153338381 residues

Word size : 16

Total number of hits satisfying chosen parameters: 10434

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCBUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Backfilltest.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	28	2.1	2943	4	US-09-404-879A-385
2	28	2.1	7152	4	US-09-167-681-29
3	28	2.1	7874	4	US-09-780-175-96
4	27	2.0	56	4	US-09-118-256-1
5	27	2.0	56	4	US-09-118-256-2
6	27	2.0	558	4	US-09-280-116-105
7	27	2.0	654	4	US-09-265-630-4
8	27	2.0	777	2	US-08-618-911-3
9	27	2.0	777	2	US-08-618-911-5
10	27	2.0	1051	4	US-09-245-041-10
11	27	2.0	1112	4	US-09-716-161A-10
12	27	2.0	1766	1	US-07-993-587A-6
13	27	2.0	1766	1	US-08-309-985-6
14	27	2.0	2028	2	US-09-211-930-12
15	27	2.0	2028	3	US-09-340-993-12
16	27	2.0	2028	4	US-09-468-442-12
17	27	2.0	3134	2	US-08-533-669A-1
18	27	2.0	3134	4	US-09-183-861-1
19	27	2.0	3134	4	US-09-022-765-1
20	27	2.0	3999	4	US-09-351-224E-9
21	26	1.9	54	2	US-08-771-624B-4
22	26	1.9	60	4	US-09-710-200-68
23	26	1.9	132	4	US-09-430-240-1
24	26	1.9	132	4	US-09-466-994-1
25	26	1.9	157	1	US-08-153-051B-27
26	26	1.9	157	1	US-08-060-952C-43
27	26	1.9	157	2	US-08-151-477A-27

28	26	1.9	157	4	US-08-464-011B-43	Sequence 43, Appl
29	26	1.9	158	3	US-08-819-867-57	Sequence 57, Appl
30	26	1.9	261	4	US-09-091-097-21	Sequence 21, Appl
31	26	1.9	267	2	US-08-924-838-3	Sequence 3, Appl1
32	26	1.9	270	2	US-08-520-678A-30	Sequence 30, Appl
33	26	1.9	270	4	US-08-897-126-30	Sequence 20, Appl
34	26	1.9	290	4	US-09-222-575-20	Sequence 25, Appl
35	26	1.9	314	1	US-08-686-878A-25	Sequence 2, Appl1
36	26	1.9	314	4	US-09-175-928-25	Sequence 25, Appl
37	26	1.9	403	4	US-09-387-212-2	Sequence 2, Appl1
38	26	1.9	403	4	US-09-948-802-2	Sequence 2, Appl1
39	26	1.9	413	4	US-09-328-111-25	Sequence 25, Appl
40	26	1.9	556	4	US-09-370-838-242	Sequence 242, App
41	26	1.9	572	4	US-09-342-653-5	Sequence 5, Appl1
42	26	1.9	722	1	US-08-698-551-5	Sequence 5, Appl1
43	26	1.9	722	2	US-08-602-228-5	Sequence 5, Appl1
44	26	1.9	722	2	US-08-649-341A-5	Sequence 5, Appl1
45	26	1.9	722	2	US-08-494-440B-5	Sequence 5, Appl1

#### ALIGNMENTS

```
RESULT 1
US-09-404-879A-385
; Sequence 385, Application US/09404879A
; Patient No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404, 879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 385
; LENGTH: 2943
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-879A-385

Query Match      2.1%; Score 28; DB 4; Length 2943;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1327 TCGGAAAAAAAAAAAAAAAAAAAAAAAAA 1354
Db 2893 TCGGAAAAAAAAAAAAAAAAAAAAAAAAA 2920

RESULT 2
US-09-167-681-29
; Sequence 29, Application US/09167681A
; Patient No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinstein, M.D., Richard M.
; APPLICANT: Kato, Thomas C.
; APPLICANT: Wood, Thomas C.
; APPLICANT: O'Brien, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167, 681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 7152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
NAME/KEY: CDS
LOCATION: (3810)...(3956)
NAME/KEY: CDS
LOCATION: (4061)...(4186)
NAME/KEY: CDS
LOCATION: (4276)...(4374)
NAME/KEY: CDS
LOCATION: (5584)...(5709)
NAME/KEY: CDS
LOCATION: (5805)...(5900)
NAME/KEY: CDS
LOCATION: (6426)...(6605)
NAME/KEY: CDS
LOCATION: (6728)...(6837)
US-09-167-681-29
```

```
Query Match
Best Local Similarity 100.0%; Score 28; DB 4; Length 7152;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1327 TCGGAAAAAAAAAAAAAAAAAAAAA 1354
DB 2373 TCGGAAAAAAAAAAAAAAAAAAAAA 2400
```

```
RESULT 3
US-09-780-175-96/C.
```

```
Sequence 96, Application US/09780175
Patent No. 6440738
GENERAL INFORMATION:
APPLICANT: ROBERT MCKAY
APPLICANT: Susan M. Freier
APPLICANT: Jacqueline Whalt
TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-BETA EXPRESSION
FILE REFERENCE: RTS-0164
CURRENT APPLICATION NUMBER: US/09/780,175
CURRENT FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 154
SEQ ID NO 96
LENGTH: 7874
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (2169)...(2297)
NAME/KEY: CDS
LOCATION: (4112)...(4150)
NAME/KEY: CDS
LOCATION: (2833)...(2913)
NAME/KEY: CDS
LOCATION: (4576)...(4614)
NAME/KEY: CDS
LOCATION: (4827)...(4994)
NAME/KEY: CDS
LOCATION: (6267)...(6431)
US-09-780-175-96
```

```
Query Match
Best Local Similarity 100.0%; Score 28; DB 4; Length 7874;
```

```
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1327 TCGGAAAAAAAAAAAAAAAAAAAAA 1354
DB 5375 TCGGAAAAAAAAAAAAAAAAAAAAA 5348
```

```
RESULT 4
US-09-118-256-1
```

```
Sequence 1, Application US/09118256
Patent No. 6238864
GENERAL INFORMATION:
APPLICANT: YAN, Lih
TITLE OF INVENTION: ANALYTE DETECTION ASSAY AND METHODS OF USE
```

```
FILE REFERENCE: 201.00010101
CURRENT APPLICATION NUMBER: US/09/118,256
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: 60/053,164
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 56
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Adaptor
US-09-118-256-1
```

```
Query Match
Best Local Similarity 100.0%; Score 27; DB 4; Length 56;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354
DB 15 CGGAAAAAAAAAAAAAAAAAAAAA 41
```

```
RESULT 5
US-09-118-256-2/C
```

```
Sequence 2, Application US/09118256
Patent No. 6238864
GENERAL INFORMATION:
APPLICANT: YAN, Lih
TITLE OF INVENTION: ANALYTE DETECTION ASSAY AND METHODS OF USE
FILE REFERENCE: 201.00010101
CURRENT APPLICATION NUMBER: US/09/118,256
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: 60/053,164
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 56
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Adaptor
US-09-118-256-2
```

```
Query Match
Best Local Similarity 100.0%; Score 27; DB 4; Length 56;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354
DB 46 CGGAAAAAAAAAAAAAAAAAAAAA 20
```

```
RESULT 6
US-09-280-116-185/C
```

```
Sequence 185, Application US/09280116A
Patent No. 6331427
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
FILE REFERENCE: 5800-24, 035800/176965
CURRENT APPLICATION NUMBER: US/09/280,116A
CURRENT FILING DATE: 1999-03-26
NUMBER OF SEQ ID NOS: 268
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 185
LENGTH: 538
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: ADAM family of metalloprotease
```

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(558)  
OTHER INFORMATION: n = a, t, c or g  
US-09-280-116-185

Query Match 2.0%; Score 27; DB 4; Length 558;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354  
DB 86 CGGAAAAAAAAAAAAAAAAAAAAA 60

RESULT 7  
US-09-265-630-4  
Sequence 4, Application US/09265630  
Patent No. 6432914

GENERAL INFORMATION:  
APPLICANT: LEVINE, BETH C  
TITLE OF INVENTION: BECLIN AND USES THEREOF  
FILE REFERENCE: 50902-D/3PW/EMW  
CURRENT APPLICATION NUMBER: US/09/265,630  
CURRENT FILING DATE: 1999-03-09  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO. 4  
LENGTH: 654  
TYPE: DNA  
ORGANISM: human  
US-09-265-630-4

Query Match 2.0%; Score 27; DB 4; Length 654;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354  
DB 626 CGGAAAAAAAAAAAAAAAAAAAAA 652

RESULT 8  
US-08-618-911-3

Sequence 3, Application US/08618911  
Patent No. 5850016  
GENERAL INFORMATION:  
APPLICANT: Jung, Rudolf  
APPLICANT: Hastings, Craig  
APPLICANT: Coughlan, Sean  
APPLICANT: Hu, David  
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN  
TITLE OF INVENTION: SEEDS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.  
STREET: 700 Capital Square, 400 Locust Street  
CITY: Des Moines  
STATE: Iowa  
COUNTRY: USA  
ZIP: 50309

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/618,911  
FILING DATE: Concurrently herewith  
ATTORNEY/AGENT INFORMATION:  
NAME: Simon, Soma  
REGISTRATION NUMBER: 37,444  
REFERENCE/DOCKET NUMBER: 365-US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (515) 248-4800  
TELEFAX: (515) 248-4844  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 777 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

FEATURE:  
NAME/KEY: CDS  
LOCATION: 28..501  
US-08-618-911-3

Query Match 2.0%; Score 27; DB 2; Length 777;  
Best Local Similarity 100.0%; Pred. No. 0.025;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354  
DB 749 CGGAAAAAAAAAAAAAAAAAAAAA 775

RESULT 9  
US-08-618-911-5  
Sequence 5, Application US/08618911  
Patent No. 5850016

GENERAL INFORMATION:  
APPLICANT: Jung, Rudolf  
APPLICANT: Hastings, Craig  
APPLICANT: Coughlan, Sean  
APPLICANT: Hu, David  
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN  
TITLE OF INVENTION: SEEDS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.  
STREET: 700 Capital Square, 400 Locust Street  
CITY: Des Moines  
STATE: Iowa  
COUNTRY: USA  
ZIP: 50309

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/618,911  
FILING DATE: Concurrently herewith  
ATTORNEY/AGENT INFORMATION:  
NAME: Simon, Soma  
REGISTRATION NUMBER: 37,444  
REFERENCE/DOCKET NUMBER: 365-US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (515) 248-4800  
TELEFAX: (515) 248-4844  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 777 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 28..501  
US-08-618-911-5

Query Match 2.0%; Score 27; DB 2; Length 777;  
Best Local Similarity 100.0%; Pred. No. 0.025;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354  
Db 749 CGGAAAAAAAAAAAAAAAAAAAAA 775

## RESULT 10

US-09-245-041-10  
; Sequence 10, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 1051  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-10

## Query Match

Best Local Similarity 2.0%; Score 27; DB 4; Length 1051;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354  
Db 7 CGGAAAAAAAAAAAAAAAAAAAAA 33

## RESULT 11

US-09-716-161A-10  
; Sequence 10, Application US/09716161A  
; Patent No. 6355482  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Susan M. Freiler  
; TITLE OF INVENTION: ANTISENSE MODULATION OF INTEGRIN BETA 4 BINDING PROTEIN EXPRESSION  
; FILE REFERENCE: RTS-0176  
; CURRENT APPLICATION NUMBER: US/09/716,161A  
; CURRENT FILING DATE: 2000-11-07  
; NUMBER OF SEQ ID NOS: 89  
; SEQ ID NO 10  
; LENGTH: 1112  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (71)...(808)  
US-09-716-161A-10

## Query Match

Best Local Similarity 2.0%; Score 27; DB 4; Length 1112;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354  
Db 1080 CGGAAAAAAAAAAAAAAAAAAAAA 1106

## RESULT 12

US-07-991-587A-6  
; Sequence 6, Application US/07991587A  
; Patent No. 5384249

## GENERAL INFORMATION:

APPLICANT: Sasaki, Katsutoshi  
APPLICANT: Watanabe, Etsuyo  
APPLICANT: Nishii, Tatsunari  
APPLICANT: Sekine, Susumu  
APPLICANT: Hanai, No. 5384249uo  
APPLICANT: Hasegawa, Mamoru  
TITLE OF INVENTION: '2 3 Stalyltransferase  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Fitzpatrick, Cella, Harper & Scinto  
STREET: 277 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10172

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44mb  
COMPUTER: IBM PC  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/991,587A  
FILING DATE: 19930526  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP-333661/1991  
APPLICATION NUMBER: JP-091044/1992  
FILING DATE: 17-12-1991  
FILING DATE: 10-04-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lawrence S. Perry  
REGISTRATION NUMBER: 31,865  
REFERENCE/DOCKET NUMBER: 1580.2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-2400  
TELEFAX: 212-758-2982  
TELEX: 236262  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1766  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: human  
CELL LINE: WM266-4 cell  
CELL TYPE: melanoma  
US-07-991-587A-6

## Query Match

Best Local Similarity 2.0%; Score 27; DB 1; Length 1766;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354  
Db 1739 CGGAAAAAAAAAAAAAAAAAAAAA 1765

## RESULT 13

US-08-309-985-6  
; Sequence 6, Application US/08309985  
; Patent No. 5494790  
; GENERAL INFORMATION:  
; APPLICANT: Sasaki, Katsutoshi  
; APPLICANT: Watanabe, Etsuyo  
; APPLICANT: Nishii, Tatsunari  
; APPLICANT: Sekine, Susumu  
; APPLICANT: Hanai, No. 5494790uo  
; APPLICANT: Hasegawa, Mamoru  
TITLE OF INVENTION: '2 3 Stalyltransferase  
NUMBER OF SEQUENCES: 7

Wed May 7 14:09:59 2003

us-10-047-825-3.rni

Page 5

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fitzpatrick, Cella, Harper & Scinto
STREET: 277 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10172
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44mb
COMPUTER: IBM PC
OPERATING SYSTEM: Dos 3.3
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/309,985
FILING DATE: 20-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/991,587
FILING DATE: 16-12-1992
APPLICATION NUMBER: JP-333661/1991
APPLICATION NUMBER: JP-091044/1992
FILING DATE: 17-12-1991
FILING DATE: 10-04-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lawrence S. Petty
REGISTRATION NUMBER: 31,865
REFERENCE/DOCKET NUMBER: 1580.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-2400
TELEFAX: 212-758-2982
TELEX: 236262
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1766
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
CELL LINE: WM266-4 cell
CELL TYPE: melanoma.
US-08-309-985-6

Query Match          2.0%; Score 27; DB 1; Length 1766;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354
Db 1739 CGGAAAAAAAAAAAAAAAAAAAAA 1765

RESULT 14
US-09-211-930-12
; Sequence 12, Application US/09211930
; Patent No. 5962265
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 5962265r1s
; APPLICANT: William Craig Moore
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; FILE REFERENCE: PHM.70296
; CURRENT APPLICATION NUMBER: US/09/211,930
; EARLIER FILING DATE: 1998-12-15
; EARLIER APPLICATION NUMBER: GB 9726851.0
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 2028
; TYPE: DNA
```

```

ORGANISM: Mus musculus
US-09-211-930-12

Query Match          2.0%; Score 27; DB 2; Length 2028;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354
Db 1975 CGGAAAAAAAAAAAAAAAAAAAAA 2001

RESULT 15
US-09-340-993-12
; Sequence 12, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228r1s
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.NI
; CURRENT APPLICATION NUMBER: US/09/340,993
; EARLIER FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-340-993-12

Query Match          2.0%; Score 27; DB 3; Length 2028;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1328 CGGAAAAAAAAAAAAAAAAAAAAA 1354
Db 1975 CGGAAAAAAAAAAAAAAAAAAAAA 2001
```

Search completed: May 3, 2003, 11:28:51  
Job time : 72 secs



GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 11:14:32 ; Search time 130 Seconds  
(Without alignments)  
12307.038 Million cell updates/sec

Title: US-10-047-825-3

Perfect score: 1354  
Sequence: 1 aattcgagcagagcttaagc.....aaaaaaaaaaaaaaaaaa 1354

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 746064 seqs, 590810554 residues

Word size: 16

Total number of hits satisfying chosen parameters: 79111

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Published Applications, NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PC1\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PC1US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	2.4	1527	10	US-09-925-297-51
2	32	2.3	1722	9	US-09-984-271-102
3	31	2.3	1844	9	US-10-001-054-23
4	31	2.3	1844	9	US-10-053-107-1
5	31	2.3	1844	9	US-10-028-072-475
6	31	2.3	1844	9	US-10-121-049-475
7	31	2.3	1844	9	US-10-123-904-475
8	31	2.3	1844	9	US-10-140-470-475
9	31	2.3	1844	9	US-10-175-746-475
10	31	2.3	1844	9	US-10-176-918-475
11	31	2.3	1844	9	US-10-176-921-475
12	31	2.3	1844	9	US-10-137-865-475
13	31	2.3	1844	9	US-10-140-474-475
14	31	2.3	1844	9	US-10-142-431-475
15	31	2.3	1844	9	US-10-143-114-475
16	31	2.3	1844	9	US-10-140-002-475
17	31	2.3	1844	9	US-10-006-836A-83
18	31	2.3	1844	9	US-10-142-419-475
19	31	2.3	1844	9	US-10-123-262-475

20	31	2.3	1844	9	US-10-142-423-475	Sequence 475, App
21	31	2.3	1844	9	US-10-006-818A-83	Sequence 83, App1
22	31	2.3	1844	9	US-10-121-050-475	Sequence 475, App
23	31	2.3	1844	9	US-10-141-755-475	Sequence 475, App
24	31	2.3	1844	9	US-10-143-032-475	Sequence 475, App
25	31	2.3	1844	9	US-10-213-145-1	Sequence 1, App1
26	31	2.3	1844	9	US-10-015-393A-83	Sequence 83, App1
27	31	2.3	1844	9	US-10-123-108-475	Sequence 475, App
28	31	2.3	1844	9	US-10-123-236-475	Sequence 475, App
29	31	2.3	1844	9	US-10-123-261-475	Sequence 475, App
30	31	2.3	1844	9	US-10-140-921-475	Sequence 475, App
31	31	2.3	1844	9	US-10-140-928-475	Sequence 475, App
32	31	2.3	1844	9	US-09-846-374-83	Sequence 83, App1
33	31	2.3	1844	9	US-10-012-121A-83	Sequence 83, App1
34	31	2.3	1844	9	US-10-015-869A-83	Sequence 83, App1
35	31	2.3	1844	9	US-10-121-045-475	Sequence 475, App
36	31	2.3	1844	9	US-10-123-292-475	Sequence 475, App
37	31	2.3	1844	9	US-10-123-903-475	Sequence 475, App
38	31	2.3	1844	9	US-10-124-819-475	Sequence 475, App
39	31	2.3	1844	9	US-10-124-822-475	Sequence 475, App
40	31	2.3	1844	9	US-10-140-925-475	Sequence 475, App
41	31	2.3	1844	9	US-10-160-498-475	Sequence 475, App
42	31	2.3	1844	9	US-10-121-041-475	Sequence 475, App
43	31	2.3	1844	9	US-10-121-047-475	Sequence 475, App
44	31	2.3	1844	9	US-10-121-047-475	Sequence 475, App
45	31	2.3	1844	9	US-10-123-215-475	Sequence 475, App

#### ALIGNMENTS

RESULT 1  
US-09-925-297-51  
Sequence 51, Application US/09925297  
Patent No. US20020081659A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA105  
CURRENT APPLICATION NUMBER: US/09/925,297  
PRIOR APPLICATION NUMBER: PCT/US00/05989  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
NUMBER OF SEQ ID NOS: 928  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 51  
LENGTH: 1527  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-297-51

Query Match 2.4%; Score 32; DB 10; Length 1527;  
Best Local Similarity 100.0%; Pred. No. 8.1e-07;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1323 TTTTCGAGAAAAA 1354  
DB 1487 TTTTCGAGAAAAA 1518

RESULT 2  
US-09-984-271-102  
Sequence 102, Application US/09984271  
Publication No. US20030040086A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins  
FILE REFERENCE: P2030P1  
CURRENT APPLICATION NUMBER: US/09/984,271  
CURRENT FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: 09/482,273

PRIOR FILING DATE: 2000-01-13  
PRIOR APPLICATION NUMBER: PCT/US99/15849  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: 60/092,921  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/092,922  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/092,956  
PRIOR FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 102  
LENGTH: 1722  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (401)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (695)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-984-271-102

Query Match 2.3%; Score 31; DB 9; Length 1722;  
Best Local Similarity 100.0%; Pred. No. 2,7e-06;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
|||||  
Db 1655 TTTTCGAAAAAAAAAAAAAAAAAAAA 1685

RESULT 3  
US-10-001-054-23  
Sequence 23, Application US/10001054  
Publication No. US20020192209a1  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Baker, Kevin  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin  
APPLICANT: Hebert, Carolyn  
APPLICANT: Henzel, William  
APPLICANT: Kabakoff, Rhona  
APPLICANT: Shelton, David  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC  
FILE REFERENCE: P3034R1PCT  
CURRENT APPLICATION NUMBER: US/10/001,054  
CURRENT FILING DATE: 2001-11-30  
PRIOR APPLICATION NUMBER: 60/059114  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/090691  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/096891

PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 60/096894  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 60/099803  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/100263  
PRIOR FILING DATE: 1998-09-14  
PRIOR APPLICATION NUMBER: 60/100390  
PRIOR FILING DATE: 1998-09-15  
PRIOR APPLICATION NUMBER: 60/101476  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/107783  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 60/108849  
PRIOR FILING DATE: 1998-11-18  
PRIOR APPLICATION NUMBER: 60/112420  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/113296  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/115554  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/115558  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/116533  
PRIOR FILING DATE: 1999-01-20  
PRIOR APPLICATION NUMBER: 60/123618  
PRIOR FILING DATE: 1999-03-10  
PRIOR APPLICATION NUMBER: 60/131294  
PRIOR FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: 60/140650  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: 60/141037  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 60/144758  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/162506  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 60/170262  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 60/187202  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 60/209832  
PRIOR FILING DATE: 2000-06-05  
PRIOR APPLICATION NUMBER: 60/232887  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: 60/180997  
PRIOR FILING DATE: 1998-11-19  
PRIOR APPLICATION NUMBER: 60/218517  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/284291  
PRIOR FILING DATE: 1999-04-12  
PRIOR APPLICATION NUMBER: 60/380137  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: 60/380138  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: 60/380913  
PRIOR FILING DATE: 1999-09-09  
PRIOR APPLICATION NUMBER: 60/403297  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: 60/423741  
PRIOR FILING DATE: 1999-11-10  
PRIOR APPLICATION NUMBER: 60/709238  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/802706  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: 60/866034  
PRIOR FILING DATE: 2001-05-25  
PRIOR APPLICATION NUMBER: 60/872035  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/882636  
PRIOR FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: 60/918585  
PRIOR FILING DATE: 2001-07-30



PRIOR APPLICATION NUMBER: 09/924419  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: 09/927796  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: 09/929404  
PRIOR FILING DATE: 2001-08-13  
PRIOR APPLICATION NUMBER: 09/941992  
PRIOR FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: 09/946374  
PRIOR FILING DATE: 2001-09-04  
PRIOR APPLICATION NUMBER: PCT/US98/18824  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: PCT/US99/00106  
PRIOR FILING DATE: 1999-01-05  
PRIOR APPLICATION NUMBER: PCT/US99/05028  
PRIOR FILING DATE: 1999-03-08  
PRIOR APPLICATION NUMBER: PCT/US99/08615  
PRIOR FILING DATE: 1999-04-20  
PRIOR APPLICATION NUMBER: PCT/US99/12252  
PRIOR FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: PCT/US99/20111  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28551  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28634  
PRIOR FILING DATE: 1999-12-01  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00376  
PRIOR FILING DATE: 2000-01-06  
PRIOR APPLICATION NUMBER: PCT/US00/03565  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: PCT/US00/04341  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: PCT/US00/04342  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: PCT/US00/05841  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: PCT/US00/06884  
PRIOR FILING DATE: 2000-03-15  
PRIOR APPLICATION NUMBER: PCT/US00/08439  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: PCT/US00/13705  
PRIOR FILING DATE: 2000-05-17  
PRIOR APPLICATION NUMBER: PCT/US00/14042  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: PCT/US00/14941  
PRIOR FILING DATE: 2000-05-30  
PRIOR APPLICATION NUMBER: PCT/US00/15264  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: PCT/US00/22031  
PRIOR FILING DATE: 2000-08-11  
PRIOR APPLICATION NUMBER: PCT/US00/23522  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: PCT/US00/30873  
PRIOR FILING DATE: 2000-11-10  
PRIOR APPLICATION NUMBER: PCT/US00/32678  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: PCT/US01/06520  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: PCT/US01/06666  
PRIOR FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: PCT/US01/17092  
PRIOR FILING DATE: 2001-05-25  
PRIOR APPLICATION NUMBER: PCT/US01/17800  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: PCT/US01/19692

PRIOR FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: PCT/US01/21066  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: PCT/US01/21735  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: PCT/US01/27099  
PRIOR FILING DATE: 2001-08-29  
NUMBER OF SEQ ID NOS: 91  
SEQ ID NO 23  
LENGTH: 1844  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-001-054-23  
Query Match 2.3%; Score 31; DB 9; Length 1844;  
Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817  
RESULT 4  
US-10-053-107-1  
Sequence 1, Application US/10053107  
Publication No. US20020192752A1  
GENERAL INFORMATION:  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune  
FILE REFERENCE: P3033R1C1  
CURRENT APPLICATION NUMBER: US/10/053,107  
CURRENT FILING DATE: 2002-01-17  
PRIOR APPLICATION NUMBER: 60/099601  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 60/107783  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 60/108802  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: 60/113296  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/131291  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/151733  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: 60/209832  
PRIOR FILING DATE: 2000-06-05  
PRIOR APPLICATION NUMBER: 60/232887  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: 09/218517  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: PCT/US99/00106  
PRIOR FILING DATE: 1999-01-05  
PRIOR APPLICATION NUMBER: PCT/US99/20111  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: PCT/US00/04342  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: PCT/US00/05601  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: PCT/US00/13705  
PRIOR FILING DATE: 2000-05-17  
PRIOR APPLICATION NUMBER: PCT/US00/14042  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: PCT/US00/14941  
PRIOR FILING DATE: 2000-05-30  
PRIOR APPLICATION NUMBER: PCT/US00/15264

```

; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/23532
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US01/17800
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/21066
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-053-107-1

Query Match          2.3%; Score 31; DB 9; Length 1844;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1817

RESULT 5
US-10-028-072-475
; Sequence 475, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
```

```

; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
```

```

PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23

PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090663
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 2.3%; Score 31; DB 9; Length 1344;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1817

RESULT 6
US-10-121-049-475
; Sequence 475, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mei-Qiang
; APPLICANT: Gerltzen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C17
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US/10/121,049
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 475
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-121-049-475

Query Match 2.3%; Score 31; DB 9; Length 1844;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAAA 1817

RESULT 7
US-10-123-904-475
; Sequence 475, Application US/10123904
; Publication No. US2003002238A1
; GENERAL INFORMATION:
```

```
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C54
CURRENT APPLICATION NUMBER: US/10/123,904
CURRENT FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 475
LENGTH: 1844
TYPE: DNA
ORGANISM: Homo Saplen
US-10-123-904-475
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 1844;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817
```

```
RESULT 8
US-10-140-470-475
Sequence 475, Application US/10140470
Publication No. US2003002231A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C160
CURRENT APPLICATION NUMBER: US/10/140,470
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 475
LENGTH: 1844
TYPE: DNA
ORGANISM: Homo Saplen
US-10-140-470-475
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 1844;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817
```

```
RESULT 9
US-10-175-746-475
Sequence 475, Application US/10175746
Publication No. US2003002270A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 475
LENGTH: 1844
TYPE: DNA
ORGANISM: Homo Saplen
US-10-175-746-475
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 1844;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817
```

```
RESULT 10
US-10-176-918-475
Sequence 475, Application US/10176918
Publication No. US2003002275A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
```

```

; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 475
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-918-475
```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 1844;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAAAAAAA 1817
```

```

RESULT 11
US-10-176-921-475
; Sequence 475, Application US/10176921
; Publication No. US20030027276A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C382
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 475
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-475
```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 1844;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAAAAAAA 1817
```

```

RESULT 12
US-10-137-865-475
; Sequence 475, Application US/10137865
; Publication No. US20030032155A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 475
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-865-475
```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 1844;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAAAAAAA 1354
Db 1787 TTTTCGAAAAAAAAAAAAAAAAAAAAAAAAA 1817
```

```

RESULT 13
US-10-140-474-475
; Sequence 475, Application US/10140474
; Publication No. US20030032156A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 475
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo Sapien
```

US-10-140-474-475

Query Match 2.3%; Score 31; DB 9; Length 1844;  
Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
|||||  
DB 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

RESULT 14

US-10-142-431-475  
; Sequence 475, Application US/10142431  
; Publication No. US20030036179A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C251  
; CURRENT APPLICATION NUMBER: US/10/142,431  
; CURRENT FILING DATE: 2002-05-10  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 475  
; LENGTH: 1844  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-142-431-475

Query Match

2.3%; Score 31; DB 9; Length 1844;  
Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
|||||  
DB 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

RESULT 15

US-10-143-114-475  
; Sequence 475, Application US/10143114  
; Publication No. US20030036180A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C211  
; CURRENT APPLICATION NUMBER: US/10/143,114  
; CURRENT FILING DATE: 2002-05-09  
; Prior Application removed - See File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 475  
; LENGTH: 1844  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-143-114-475

Query Match

2.3%; Score 31; DB 9; Length 1844;  
Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1324 TTTTCGAAAAAAAAAAAAAAAAAAAA 1354  
|||||  
DB 1787 TTTTCGAAAAAAAAAAAAAAAAAAAA 1817

Search completed: May 3, 2003, 12:17:40  
Job time: 149 secs



Journal  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (15-JUL-1994) Annette H. Ross, University of Queensland,  
Botany, St. Lucia, Brisbane, Queensland, 4072, Australia

FEATURES  
source  
1. 1335  
/organism="Cenchrus ciliaris"  
/cultivar="Billoela"  
/db\_xref="taxon:35872"  
/clone="px7"  
/issue\_type="wounded and unwounded leaves and stems"  
/clone\_id="lambda gt-10"  
1. 57  
58. 981  
/EC\_number="1.11.1.7"  
/note="expressed in leaves and is wound-induced"  
/product="peroxidase"  
/codon\_start=1  
/protein\_id="AA020472.1"  
/db\_xref="GI:520568"  
/translation="MAAPTMOCILATISLSTFAHOLSTFFYASSCPNLOTVVRAM  
TOAYSSPPMGASTIRLFFHDFVCGDASITLRSRSGKSGAGNANVIRGYIDITIK  
KNVACAPGVYSCATYPLAARCPNIGSGPTNVPGRDSTMANSLANONLPPT  
SLGTLISLFGRLSARDMIALSGAHQAQCTTFRGRIYDNTIDASFDALOOOTCP  
SGGDNLPAPIDAOIPARFDNAYTNLVRGLHSDLEFNSSDPLVNOISSFSQ  
FNSDFVAMIKMGINAGNOVRNCRVNS"  
58. 126  
/note="putative"  
127. 978  
/product="peroxidase"  
982. 1335  
1335  
/note="39 A residues"  
BASE COUNT 264 a 440 c 374 g 257 t  
ORIGIN  
Query Match 49.6%; Score 671.4; DB 8; Length 1335;  
Best Local Similarity 81.5%; Pred. No. 1.5e-87;  
Matches 832; Conservative 0; Mismatches 171; Indels 18; Gaps 4;

55 CAGACAGCTGGCCGCTCTCCACCTTGATGCAATGCTGGTCCGCTTCTCTC 114  
|||||  
46 CAGACAGCTGGCCGCTCTCTACCTTGATGCAATGCTGGTCCGCTTCTCTC 105  
|||||  
115 TCCTGTGTCGCCACGACAGCTTCTGCGCCACCTTCTCTCTCTCTCTCTCT 174  
|||||  
106 TCCTGTGTCGCCACGACAGCTTCTGCGCCACCTTCTCTCTCTCTCTCTCT 165  
|||||  
175 CAGACAGCTGGCCGCTCTCCACCTTGATGCAATGCTGGTCCGCTTCTCTC 234  
|||||  
166 CAGACAGCTGGCCGCTCTCTACCTTGATGCAATGCTGGTCCGCTTCTCTC 225  
|||||  
235 TCCTGTGTCGCCACGACAGCTTCTGCGCCACCTTCTCTCTCTCTCTCTCT 294  
|||||  
226 TCCTGTGTCGCCACGACAGCTTCTGCGCCACCTTCTCTCTCTCTCTCTCT 285  
|||||  
295 CTTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354  
|||||  
286 TCCTGTGTCGCCACGACAGCTTCTGCGCCACCTTCTCTCTCTCTCTCTCT 345  
|||||  
355 GTCATGACACATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 414  
|||||  
346 GTCATGACACATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 405  
|||||  
415 GACATCTGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 474  
|||||  
406 ACATGCTGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 465  
|||||  
475 GTGCGCGCTGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 534  
|||||

Db 466 GTGCGCGCTGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 525  
Oy 535 CCGCCCG 594  
Db 526 CCGCCCG 579  
Oy 595 CCG 654  
Db 580 GCG 636  
Oy 655 TTCG 714  
Db 637 TTCG 696  
Oy 715 CAGACAGCTGGCCGCTCTCCACCTTGATGCAATGCTGGTCCGCTTCTCTCTCT 774  
Db 697 CAGACAGCTGGCCGCTCTCCACCTTGATGCAATGCTGGTCCGCTTCTCTCTCT 756  
Oy 775 GTGAGCTGACACGCGCGCTTCTCCACCACTGCTGCGCGCGCGCGCGCGCGCG 834  
Db 757 GCGAGGTTGACACGCGCGCTTCTCCACCACTGCTGCGCGCGCGCGCGCGCGCG 816  
Oy 835 GACACAGAGCTTCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 894  
Db 817 GACACAGAGCTTCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 876  
Oy 895 GCTGCGCTTCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 954  
Db 877 CCGTCCGAGTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 933  
Oy 955 CTTACCG 1014  
Db 934 -----GCCAACG 987  
Oy 1015 CGATGCGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1074  
Db 988 CAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1047  
Oy 1075 C 1075  
Db 1048 C 1048

RESULT 2  
AP005613/c 160562 bp DNA linear HTG 02-AUG-2002  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) chromosome 2 clone  
OSJNB0082C09, \*\*\* SEQUENCING IN PROGRESS \*\*\*; in ordered pieces.  
AP005613  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
HTG; HTGS\_PHASE2.  
ORGANISM  
Oryza sativa (japonica cultivar-group) (cultivar: Nipponbare) DNA,  
clone: OSJNB0082C09.  
Oryza sativa (japonica cultivar-group); Embryophyta; Tracheophyta;  
Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae;  
Eriophytidae; Oryzaceae; Oryza.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Direct Submission  
Submitted (01-AUG-2002) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program; Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/  
Tel: 81-298-38-7441, Fax: 81-298-38-7468)  
NOTE: It currently consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces is believed  
to be correct as given, however the sizes of the gaps between them









OY 446 CCAACTTCGCGCGCCGACGCTGCGCTGGCGGCGGAGCTGACGACGG 505  
 Db 362 TCTATTGAGGCGGGCCACATGACGCTCCACTAGGCGAGCGCGGAGG 421  
 OY 506 CCAAGCGCTGCTGCGCAACGCCCGCGCGCGGCGGCGGCGGCGGCGG 565  
 Db 422 CGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 481  
 OY 566 TCTCCTGTTGCGGAGCGGCGCTGCTGCGGCGGCGGCGGCGGCGG 625  
 Db 482 CCAACTGTTGCGGAGCGGCGCTGCTGCGGCGGCGGCGGCGGCGGCGG 541  
 OY 626 ACACGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 685  
 Db 542 ACACGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601  
 OY 686 TCAACGCTTCCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 745  
 Db 602 TCGACCGCGGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 661  
 OY 746 ACCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 805  
 Db 662 ACCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 718  
 OY 806 TCTGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 865  
 Db 719 TGTGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 778  
 OY 866 ACCGCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 925  
 Db 779 ACCGCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 838  
 OY 926 CCATGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 985  
 Db 839 CTAATGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 898  
 OY 986 ACTGCGCGG 993  
 Db 899 ACTGCGAGG 906

RESULT 6  
 AP003991/c 149417 bp DNA linear HTG 21-MAR-2002  
 LOCUS  
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone  
 OJ1077\_A12, \*\*\* SEQUENCING IN PROGRESS \*\*\* in ordered pieces.  
 ACCESSION AP003991  
 VERSION AP003991.1 GI:15076799  
 KEYWORDS HTG, HTGS, PHASE2.  
 SOURCE Oryza sativa (japonica cultivar-group) (cultivar: Nipponbare) DNA,  
 clone: OJ1077\_A12  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1  
 REFERENCE  
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.  
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC  
 clone: OJ1077\_A12  
 JOURNAL Published Only in Database (2001)  
 2 (bases 1 to 149417)  
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-AUG-2001) Takuji Sasaki, National Institute of  
 Agrobiological Resources, Rice Genome Research Program, Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgrp.dna.affrc.go.jp/,  
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)  
 The nucleotide sequence of this BAC clone was generated by  
 combining Monsanto and RGP-Japan sequencing data.  
 NOTE: It currently consists of 1 contigs. Gaps between the contigs  
 are represented as runs of N. The order of the pieces is believed  
 to be correct as given, however the sizes of the gaps between them

are based on estimates that have provided by the submitter. This  
 sequence will be replaced by the finished sequence as soon as it is  
 available and the accession number will be preserved.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

FEATURES  
 source  
 1. 149417  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="2"  
 /clone="OJ1077\_A12"

BASE COUNT 41306 a 33053 c 33283 g 41528 t 247 others  
 ORIGIN

Query Match 26.9%; Score 364.6; DB 2; Length 149417;  
 Best local Similarity 68.0%; Pred. No. 1.5e-43;  
 Matches 615; Conservative 0; Mismatches 249; Indels 41; Gaps 6;

OY 137 TCTGCGCCACGTTCTATGCGCTCTCTGCCCCCAACCTGACAGACATGCGGCGGCA 196  
 Db 43397 TGTGCGAGGCTACTATGCTAAGAGCTGCGCGCGGCGGAGTCTGCTGCTGCTGA 43338  
 OY 197 TGACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 256  
 Db 43337 TGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCTTCTTC 43278  
 OY 257 ACAGCTGCTTCTTGAAGGCTGCGAGCGGATGATCTTCTGACGCGG----- 304  
 Db 43277 ACAGCTGCTTCTTGAAGGCTGCGAGCGGATGATCTTCTGAGAGCGGCGGCGG 43218  
 OY 305 ---GAGGGGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 361  
 Db 43217 TCACCGGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 43158  
 OY 362 ACACGATCAAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 421  
 Db 43157 ACAGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 43098  
 OY 422 TCGCGCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 481  
 Db 43097 TCGCGCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 43038  
 OY 482 TCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541  
 Db 43037 TGGAGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 42978  
 OY 542 CGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601  
 Db 42977 CG-GTGCAGGCTTCACTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 42920  
 OY 602 ACATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 661  
 Db 42919 ACATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 42860  
 OY 662 GCGGCGATCA---CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 718  
 Db 42859 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 42800  
 OY 719 CGTGCG---CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 775  
 Db 42799 TGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 42740  
 OY 776 TGAGCTTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 835  
 Db 42739 ACGTGTGAGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 42680  
 OY 836 ACCAGAGGCTTCAAGCGGCGG-----GTGCGAGAGAGCGGCGGCGGCGG 877  
 Db 42679 ACCAGAGGCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 42620



TITLE	26.8%; Score 363.4; DB 8; Length 1306;
JOURNAL	Best Local Similarity 63.5%; Pred. No. 3,7e-43;
REFERENCE	Matches 591; Conservative 0; Mismatches 331; Indels 9; Gaps 2;
AUTHORS	
TITLE	
JOURNAL	
FEATURES	
source	
gene	
CDS	
gene	
CDS	
BASE COUNT	
ORIGIN	
Query Match	
Best Local Similarity	
Matches 591; Conservative	
0; Mismatches 331; Indels 9; Gaps 2;	
94	
80	
154	
140	
214	
200	
274	
260	
334	
317	
394	
377	
454	
437	
514	

D6	497	GCCCTGGCAAAATACGAGCCTCCCTGCCCTTCTCTAAGCCATCGAGAACTTATCGGCATP	556
OY	574	TTCGGCAGAGCAGAGGCGCTGTGCGCGCGGCAGCATNMGAGCGGCTGCGGGCGGCACACATC	633
D6	557	TTCCTCAGAAAAGGACTTCGACGCAACGCACTGGTTGCTCTTAGAGAGCACACAGATC	616
OY	634	GGGCAGAGCCCGGTGCACCACTTCCGCGCGCATTAAGCGGACACCAGATCAAAGCC	693
D6	617	GGGAGAGCGCAGTGCCAGAACTTCAGGGAGACAGATCTTACAACGAGACCAATCGATWTC	676
OY	694	TCCCTGCGCGCGCGTGGCGGAGCAGACGTGCCCGGGGTC-----CGCGCGCGAGCGCAAC	747
D6	677	GCCCTGCAGCGAGAACGCCGCACTGCCCCAGCGCCGAGCGGCAACGGCGCAGACGA	736
OY	748	CTGCGCGCCCATCGACCTGCGAGACGCCGGTGAAGTTTCGACACGGCTACTTACCAACTCG	807
D6	737	CTGCGCGCGCTGAGACGACGACGCCCAACGCCCTTCAGCAAAAGCTCTCTACACAACTG	796
OY	808	CTGCGCGCGCGCGCGCTGTTCCACTGCGAGCAGGAGGACTTCMAAGCGCGGGGTGCAGAGC	867
D6	797	CTGTGCACAAAGAGGGCTCTCGACACTTCGACACAGAGTGCTTTCAACGCGCGGCAAGCGG	856
OY	868	GCGCTGTGAGCAGTACAGCCGCAAGCCCTCTGCTTTCAAGCGCCACTTCGTGCGAGCC	927
D6	857	AACACGCTGCGGAACTTCGCTCCAAACGCGCGCGGTTCAGCAGCCGCTTCACGAGAGCC	916
OY	928	ATGATTGATGTGGGCAACGTTGGGGTGTGTCACCGGCAACCGCGCGAGATCAGCGCAAC	987
D6	917	ATGTGTAAGAATGGGGAACAATCTCGCGCGCTGACCGGAGACGAGGCGAGATCAAGCTCAGC	976
OY	988	TGCCGGGTCTCAACAGCTAATACGACGCA	1018
D6	977	TGCTCCAGAGGTGAACCTAATTAAGAGGTA	1007
RESULT 9			
LOCUS	CCU12315	1255 bp	mRNA linear PLN 02-AUG-1994
DEFINITION	Cenchrus ciliaris clone PX18 peroxidase mRNA, complete cds.		
VERSION	012315		
KEYWORDS	GI:520569		
SOURCE	Cenchrus ciliaris.		
ORGANISM	Cenchrus ciliaris. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Paniceae; Cenchrus. 1 (bases 1 to 1255)		
REFERENCE	Ross,A.H. Investigation of peroxidase genes and genetic transformation in buffel grass		
AUTHORS	JOURNAL Thesis (1994) Botany, University of Queensland		
REFERENCE	AUTHORS		
TITLE	Direct Submission		
JOURNAL	Submitted (15-JUL-1994) Annette H. Ross, University of Queensland, Botany, St. Lucia, Brisbane, Queensland, 4072, Australia		
FEATURES	Location/Qualifiers		
source	* . organism="Cenchrus ciliaris" cultivar="Biloela" db_xref="taxon:35872" tissue_type="wounded and unwounded leaves and stems" clone_lib="lambda gt-10" 1..64 65..1006 EC_number="1.11.1.7" note="expressed in leaves and is not wound-induced" codon_start=1 product="peroxidase" protein_id="AAZ0473.1" db_xref="GI:520570" translation="MASSVSGLLIMLCMAVAASAQLSATFYDTSCPNNLSTIKSAVTAA		
.5' UTR			
CDS			





[illegible]

RESULT 11	LOCUS	DEFINITION	VERSION	KEYWORDS
AP003544/c	135792 bp	DNA	linear	HTNG 21-MAR-2002
AP003544	Oryza sativa (japonica cultivar-group)	chromosome 6	clone P0599C12,	
*** SEQUENCING IN PROGRESS ***		in ordered pieces.		
AP003544				
AP003544.1	GI:13810550			
HTNG: HTNG_PHASE2.				

SOURCE	Oryza sativa (Japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0599C12.
ORGANISM	Oryza sativa (Japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzeae; Oryza.
REFERENCE AUTHORS TITLE	1 Sasaki,T., Matsumoto,T. and Yamamoto,K. Oryza sativa nipponbare(GA) genomic DNA, chromosome 6, PAC clone:P0599C12 Published Only In Database (2001).
JOURNAL	2 (bases 1 to 135792)
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Direct Submission
JOURNAL	Submitted (25-Apr-2001) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsusaka@affrc.go.jp, URL:http://rgrp.dna.affrc.go.jp/ Tel:+81-298-38-7441, Fax:+81-298-38-7468)
COMMENT	NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved. * NOTE: This is a "working draft" sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved.
FEATURES	Location/Qualifiers
source	1..135792 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="6"
BASE COUNT	37908 a 30304 c 30206 g 37124 t 250 others
ORIGIN	
Query Match	26.0%; Score 352.2; DB 2; Length 135792;
Best Local Similarity	77.6%; Pred.No. 9.3e+42;
Matches 426; Conservative 0; Mismatches 123; Indels 0; Gaps 0;	
OY 454	CTCGGCGGGCCGACGTGAGCGGCCGGCGGCGGCGGCAGACGAGCGCACGGCC 513
DB 25700	CTGGTGTGGCCAAGTGAGACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 25641
OY 514	TTCGCTGCCAACAGCAACCCTCCCGCCCCGAGCGGCGACGCTCGACGCTATCTCC 573
DB 25640	AGCGCGCGGAACAGCAACCTCGCGGGGGCCCGGGGTGAGACTTCGCGACAGCTCATCAGATTG 25581
OY 574	TTGGGCAAGGAGGGCGCTGCGCGCGCGGACATGAGGGCGGTGGGGGGGGGACACATC 633
DB 25580	TTTCGGCAACAAGGGCTCTTGGGGGGCGACATGAGGGCGGTGGGGGGGGCCACACCATC 25521
OY 634	GAGGAGGCGCGGTGACCAACCTTCGCGGGCGGCATCTACGGCGACACCGATCAAGCCC 693
DB 25520	GCGCGCGCCAGATGCGAGTTCTTCGCGACGCCCATCTACCGAGCGCAACATCAAGCCG 25461
OY 694	TTCCTTCGCGCGGTGGGGGAGCAAGATGCGCGCGGTCCGGCGCGCGAGCGCAACTGGGG 753
DB 25460	TCTCTTCGCGCTTCGGGAGCAAGAGTGCCTCCGCGCGCGCGCGCGCGCAACTTGGG 25401
OY 754	CCCATCCAGCTGAGAGCGCGGTGAGGTTCGACAGGCGCTACTTCACCAACTGCTGTGCG 813
DB 25400	CCGTTGAGCTGAGAGCGCGCGCGCTTCGACAAAGCGCTACTACGAAAGCTCGTGTGCG 25341
OY 814	CAGGAGGGGCGCTTTCACCTCGAGACCAAGAGCTCTTCAACGGCGGGGTGGCAAGCGCGTG 873
DB 25340	CAGGCGGGCGCTGCTCAGCTCGACCAAGGAGGCTCTTCAATGGCGGCTGCGAGAGGCGCTTC 25281
OY 874	GTGAGGAGTAGACGCGCGCGCTCGCTCTTCAACGCGCGAGCTTCGTGGAGGCATGATT 933







[illegible]

JOURNAL	sequences
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 1360)
TITLE	Padejimas L.S. and Relchert,N.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (06-DEC-1997) Department of Plant and Soil Sciences,
AUTHORS	Mississippi State University, Box 9555, Mississippi State, MS
TITLE	3 (bases 1 to 1360)
JOURNAL	Padejimas,L.S. and Relchert,N.A.
REFERENCE	Submitted (25-JUL-2001) Department of Plant and Soil Sciences,
AUTHORS	Mississippi State University, Box 9555, Mississippi State, MS
TITLE	Sequence update by submitter
JOURNAL	On Jul 25, 2001 this sequence version replaced gi.393232.
REMARK	Location/Qualifiers
COMMENT	1..1360
FEATURES	/organism="Zea mays"
SOURCE	/db_xref="taxon:4577"
CDS	72..1043
	/codon_start=1
	/product="putative peroxidase P7X"
	/protein_id="AACJ954.2"
	/db_xref="gi:15011986"
	/translation="MAASVASCILSHSLAVYLVALSAASAQSLSTFYDRSCPNNL STRGNVAAYROPRVSGASILRLHPHDCVRKCDASILNPTSEOGGPNLILNPPR GVVVNSIKAOVESVCPIVSADILAVARODVYVGSPWVLIGRDSTASFPOG TSDPEPTSSIRLLISAVNRKNLNPTDMVALSCAHITGOFGCSFNHDHYNDTNISA FASLRANCPRAGSTALAPLDITTPRNPFNAATYTNILSKGLIHSDELFPNGSTDST VASPSAOSHVLQQRRLRHGHODGPDPDNPADDAQLLEQOLVNT"
BASE COUNT	303 a 420 c 370 g 267 t
ORIGIN	
Query Match	24.7% Score 334.6; DB 8; Length 1360;
Best Local Similarity	64.4%; Pred. No. 5.2e-33;
Matches 583; Conservative	0; Mismatches 309; Indels 13; Gaps 5;
Oy	112 CTCGCCGTGTCCACCACAGCATCTCGGCCACAGTCCTTGTCCTCCTCCCAC 171
Db	138 CTGGCGTCGGCGGCCTCGCGCAGCTGTCTGTCGACGTTCTACGACAGTCATCCCCAAC 197
Oy	172 CTCGACAGCATCTTGTGGCGGCGCATGACCCAGGCCGTCGCACACTGACAGCAAGATGGCG 231
Db	198 GCGCTGTCCACATYCAGAGGAGCGGCTGAATCCGCGGTGAAGCAGGACCTCCGCTGGGG 257
Oy	232 GCCTCTGTGTCAGGGCTCTTCTTCACGACAGTCGTCGTTCAAAGGCTCGACGGATGATC 291
Db	258 GCGTCGCTGTCAAGCTCATTTCCACGACATCTTTTGTCCGGGGCTCGACGGGCTCCTT 317
Oy	292 CTTCT--CGAGCGCGGAGGAGAAGACCGCGGGCCGAACCTGAACTCGTGC---GC 345
Db	318 CTGCTGAACACACGTCAGAGGAGAGCACCAAGGGCCGGAATCTTAATCTGAAACCCAAG 377
Oy	346 GGCTTAGGTCAATCACAACATCAAGAGGGAACGTCGAGGGCGGATCCCCCGCGCTCGTG 405
Db	378 GGCTTCGTTTGTCTGAACACGATCAAGGCGCAGTGGAGTCCGTGTCCCGGGGATCTGTC 437
Oy	406 TCGTGCCGCACATCTTCGCGCTTGCACGCGCGACGAGCAAGAACCAACTTCTCGGCGGGCG 465
Db	438 TCCGCGCCGACATCTTCGCGCTTGCACGCGCGACGAGCAAGAGTGTATCGCTCGGCGGCT 497
Oy	466 ACCGTGAGGCGTGGCGGTGGGGGGGGGAGACTGACAGCAAGCGCAGCGGCTCGCTCGCACAC 525
Db	498 TCCTGACACATTTCGTTAGGGCGAAGGAGACTCTACCGCTTCATTTCCAGAC---AGATA 554
Oy	526 AGCAACCCCGCCCGGACGCGACGCGCTGCGGACGCTCATCTCCCTGTTCGGCAGGACG 585

Db 555 AGGACCCCCCAGCTCCGACGCTTAGCCTCCGACAGCTTTGTGTGCTATACAAAGAAG 614  
QY 586 GGGCTGTCCCGCGGACATGACGGCGCTGTGCGGGCGCGACACCATCGGGCAGGCCGG 645  
Db 615 AATCTAACCCAAACGACATGGTTGACTCTCAGAGAGCTCACAGATCGGACAGGCGAG 674  
QY 646 TGCACCACTTCGCGGGCGCATCTACGGCGACACCGACATCAACGGCTCTCGCGGGCG 705  
Db 675 TCGTCAGAGTTCAAGACCAATCTACAAAGACACCAACATCAACTCGGCTTCGGGGCG 734  
QY 706 CTGCGGACACAGAGTGGCCCGGGTCCGGCGCGGCGAACCTGGGGCCCATCGACGTG 765  
Db 735 TCGCTCAGGGCCAACTGCCCCAGGGGAGGACACCG--CCCTTGGCCGCTGGACACC 791  
QY 766 CAGACGCGGTGAGGTTCGACACGGGCTACTTACCAACCTGTGTGCGGGCGGGCTG 825  
Db 792 ACGACGCCCAAGCGGCTTGACACAGCCCTACTACACCAACTGTGTGCCAGAAAGGGGCTC 851  
QY 826 TTCCACTCGGACAGAGAGCTTTCAACGGCGGGTGGCAGAGCGGCTGTGAGGCAGTAC 885  
Db 852 CTGCACCTCGGACAGAGAGCTTTCAACAGGCGGACACCGACAGCAGGTCAGAGAGCTTC 911  
QY 886 -AGCGCCAGCGCTCGCTTTCAACGCCACTTCGTGGCAGCCATGATTAAGATGGCAA 944  
Db 912 GCGTCCACAGACGCTCGGCTTTCAACAGGCGCTTCGCCACGGCCATGGTCAAGATGGCAA 971  
QY 945 CGTTGGGGTGTCTACCGGACCGCGCGACAGATCAAGGCGCAACTGCCGGGTGCTCAACAG 1004  
Db 972 CCTCAGCCCCCAGACCGGAAACCGAGGGGAGATCAGGGGCGAGTGTGAAGGTCAACTC 1031  
QY 1005 CTAGA 1009  
Db 1032 GTAAA 1036

Search completed: May 3, 2003, 09:09:05  
Job time : 3147 secs

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 06:47:20 ; Search time 241 Seconds  
(without alignments)  
12652.306 Million cell updates/sec

Title: US-10-047-825-3

Perfect score: 1354  
Sequence: 1 aattcgccagcagcttaagc.....aaaaaaaaaaaaaaaa 1354

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	414.2	30.6	1156	22	AAH44077
2	401	29.6	1171	22	AAH44082
3	360.8	26.6	1310	22	AAH44085
4	356	26.3	1223	22	AAH44083
5	345.6	25.5	1295	24	ABN87249
6	338.6	25.0	1158	22	AAH44080
7	334.6	24.7	1379	22	AAH44081
8	314.4	23.2	1230	22	AAH44081
9	287.2	21.2	1171	21	AAH44081

10	274.4	20.3	1224	21	AAH44087	Pinus radiata pero
11	273	20.2	1250	22	AAH44088	Oryza sativa pero
12	268.6	19.8	801	21	AAH44089	Pinus radiata pero
13	258.6	19.1	1391	21	AAH44091	Eucalyptus grandis
14	254.4	18.8	975	22	AAH44092	Synthetic Nicotian
15	245	18.1	975	22	AAH44093	Synthetic Nicotian
16	245	18.1	1218	21	AAH44094	Arabidopsis thalia
17	244.6	18.1	1444	22	AAH44095	Oryza sativa pero
18	243.4	18.0	1220	21	AAH44096	Arabidopsis thalia
19	240	17.7	1348	22	AAH44097	Rice green leaf pe
20	238.8	17.6	1233	22	AAH44098	Oryza sativa pero
21	236.2	17.4	1218	22	AAH44099	Oryza sativa pero
22	234	17.3	916	22	AAH44100	Oryza sativa pero
23	233.6	17.3	1348	22	AAH44101	Oryza sativa pero
24	231.4	17.1	1137	22	AAH44102	Oryza sativa pero
25	226.6	16.7	1355	22	AAH44103	Maize per5 root pr
26	220	16.2	1325	22	AAH44104	Oryza sativa pero
27	216.8	16.0	1317	22	AAH44105	Oryza sativa pero
28	213.4	15.8	1256	22	AAH44106	Oryza sativa pero
29	210.6	15.6	1306	22	AAH44107	Oryza sativa pero
30	210.2	15.5	1144	18	AAH44108	Oryza sativa pero
31	190.2	14.0	1370	22	AAH44109	Stylosanthes humil
32	190.2	14.0	1445	22	AAH44110	Oryza sativa pero
33	186.4	13.8	1433	22	AAH44111	Oryza sativa pero
34	184.6	13.6	790	24	ABN87205	Oryza sativa pero
35	184.4	13.6	787	24	ABN87206	Oryza sativa pero
36	184.4	13.6	792	24	ABN87207	Oryza sativa pero
37	182.8	13.5	586	21	AAH44112	Oryza sativa pero
38	182.8	13.5	1204	22	AAH44113	Oryza sativa pero
39	182.8	13.5	646	21	AAH44114	Oryza sativa pero
40	174.2	12.9	1244	21	AAH44115	Arabidopsis thalia
41	174	12.9	938	21	AAH44116	Arabidopsis thalia
42	172.6	12.7	1247	21	AAH44117	Arabidopsis thalia
43	171.8	12.7	1014	21	AAH44118	Arabidopsis thalia
44	171.4	12.7	1073	21	AAH44119	Arabidopsis thalia
45	169.2	12.5	1186	21	AAH44120	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1	AAH44077	standard; cDNA; 1156 BP.
AC	AAH44077	
AC	AAH44077	
DT	12-SEP-2001	(first entry)
DE	Oryza sativa peroxidase s4235 encoding cDNA SEQ ID NO:13.	
KW	Oryza sativa; rice; peroxidase; POX; characteristic: gene expression; modification; plant; bacterial infection; Magnaporthe grisea; ss.	
KW	Oryza sativa.	
OS	Oryza sativa.	
XX	Key	Location/Qualifiers
FT	CDS	75..1058
FT	/*tag= a	/product= "peroxidase s4235"
XX	WO200142475-A1.	
XX	14-JUN-2001.	
XX	08-DEC-2000; 2000MO-JR08728.	
XX	10-DEC-1999; 99JP-0352472.	
XX	(NORQ ) JAPAN MIN AGRIC FORESTRY & FISHERIES.	
XX	Onishi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;	
XX	Hiraga S;	











XX 28-SEP-2001; 2001WO-AU01221.  
PF  
FR XX  
XX 29-SEP-2000; 2000AU-0000419.  
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
PP (AGRE-) AGRESEARCH LTD.  
XX Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;  
PI WPI: 2002-444025/47.  
P-P-PSDB; ABB78992.

XX Novel nucleic acid encoding lignification and cellulase enzymes or  
PT their related enzymes useful for modifying lignin biosynthesis and  
PT cellulose degradation in plants to manipulate plant cell wall

PS Claim 8; Fig 105; 436pp; English.

CC The present invention describes a nucleic acid (I) or its fragment  
CC encoding caffeoyl-CoA 3-O-methyltransferase (COMMT), cinnamyl alcohol  
CC dehydrogenase (CAD), caffeic acid O-methyltransferase (OMT), peroxidase  
CC cinnamate-4-hydroxylase (C4H), cinnamoyl-CoA reductase (CCR), peroxidase  
CC (PER), cellulase (CEL), ferulate-5-hydroxylase (F5H), phenyllalanine  
CC ammonia lyase (PAL) or 4-commarute-CoA ligase (4CL) from perennial  
CC sequence (lotium perenne) or fescue species. (II), its nucleotide  
CC sequence information and/or single nucleotide polymorphisms is useful as  
CC a molecular genetic marker (I) can be used for modifying lignin  
CC biosynthesis and/or cellulose degradation in a plant to manipulate cell  
CC walls. (I) or its fragments are useful for isolating cDNAs and genes  
CC encoding homologous proteins from the same or other plant species, as  
CC hybridisation probes to screen libraries from the desired plant. Short  
CC segments of (I) or its fragment are useful in amplification protocols  
CC to amplify longer nucleic acids or its fragments encoding homologous  
CC genes from DNA or RNA. (I) or its fragments are useful as molecular  
CC genetic markers for quantitative trait loci (QTL) tagging, QTL mapping,  
CC DNA fingerprinting, and in marker assisted selection, particularly in  
CC ryegrass and fescues, and in forage and turf grass improvement, e.g.  
CC tagging QTLs for herbage quality traits, dry matter digestibility,  
CC mechanical stress tolerance, disease resistance, insect pest resistance,  
CC plant stature, leaf and stem colour. The present sequence represents a  
CC lotium perenne (perennial ryegrass) nucleotide sequence from the present  
CC invention.

XX  
XX  
SO Sequence 1295 BP; 295 A; 420 C; 337 G; 241 T; 2 other;

Query Match 25.5%; Score 345.6; DB 24; Length 1295;  
Best Local Similarity 62.7%; Pred. No. 2.6e-57;  
Matches 591; Conservative 0; Mismatches 339; Indels 12; Gaps 3

G 64 GCATGCGCTGCCACCTGTATGCATACTCCGGTGCGCGTTTCCTCTCTCTGTGTC 123  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
D 85 GCCATTGCATCTGCTCTTGCACTTTCTTTGGTTGGTTGGCCCTGGCAGCAGGCG 144  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Y 124 GCCCACGACAAGCTCTGCCACAGTTCTATGCTGCTCTCCCCAACCTGACAGAGATC 183  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
D 145 GCGTGGCGGACGCTGCTGTGACGTTCTACGACAGTGTGCTGCCACAGAGCCCTTGACC 204  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Y 184 GTTCGGGCGGCGATGACCCAGGCGCTGCGCATGTGACAGAGAGATGGCGCTCTGTCTC 243  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
D 205 ATCAAAGAGCGGCGGACGCGCGCGTAGTACAACCCCACCATAGGCGGCGCTCTGCTC 264  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Y 244 AGGCTCTTCTTCCACAGCTCTGCTTCTTCAAGGCTGCGACGAGATGAGTCTTCTGACAGCC 303  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
D 265 AGGCTGCACTTCCACGACTGCTGTGTATTGAAGCTGTAGCG---CGTCCGTTTCTCTGTG 321  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Y 304 GGAGGGAGAAGAACCGCGCGGCGGCAACTGAACTGCTGTGCGCGGCTTTGAAGTCAATGAC 363  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
D 322 GGCAACACGACGAGAACCGGCTGCGAACGCGCGGCTGTGCTGTGGCTTCTGACATGAC 381  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Y 364 ACCATTAAGGGAAGACGTGCGAGGCGGCGTCCCGCGCGTGTGCTGCGCCGACATCTCTC 423  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
D 382 AACATCAAGACCCAGCTCTGAGGGTATCTCTGACAGACAGACCTCTCTCTGCGCGACATCTCTC 441  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

424	GGCGTTCCGCGCGCGACGAGAACCAACTTTCTGGCGGGCGGACCTGGAGAGCGGCTC	483
442	ACCGTCGCGCCCGCGGACCTCTGTCTGTGCCCTGGCGGGCGCTCATGAGACGTTCCTCA	501
484	GGGCGGCGGACTCGACGAGAGGGCGCGCTCGCTGCGCAACGACACCCCGCCCGC	543
502	GGCAGAGCGGGACTCGAGC-----AGCGCAACGGGGAATTCGGGGACCTCCGGGCGCTT	555
544	ACGCGCAGCTCTCGGACAGCTCATCTCCCTTTTGGGCAAGGAGGGCCCTGTCCCGCGCGAC	603
556	GGCTCCAGCCCTCGCAGCTCCAGGCGCTCGTTGGCCAAAGAACTCAACACACTTGGAC	615
604	ATGACGGGCGGTGTGGGCGGCGACACCATGGGCGAGCGCGGTGGCAACACTTCCGCGG	663
616	ATGGTGGCGCTCTCGGCGCGCACCATGTGGCGGGCGCAGTGGCCAGAACTTCGGGTG	675
664	CGCATCTA--CGCGCAGACCGACATCAACGCGCTCTTTCGGCGGCGGTGGCGGCAAGAC	720
676	CGGATCTACGGCGGCGCGACACCAATCAAGCGCCGCTTCGCGACCTTCGCTCAAGGGCCAA	735
721	TGCCCGCGGTCCGGCGGCGAGAGGCAACCTGGCGCCATTCGACGTGGCAGAGGCCGGGTGAG	780
736	TGCCCGCGGTCCGGCGGCGACAGGCAACCTGGCGGCGGTGAGCGCGACAGCGCGCAACGCC	795
781	TTTGCACAGGGCTCTTTCACCAACCTGTGTGCGGGGGGCGCTGTTCACCTCGGACAG	840
796	TTTGCACAGCGCTCTTACACCGACGACTGTCTCTCCAGAAAGGGGCTCTCTCATTTCCGACAG	855
841	GAGCTCTTCAACGGCGGGGTGCGAGGACGCGCTGTGTGAGGCGATTCACAGCGCAGCGCTCG	900
856	GTCGCTTTCACAAACAGGCGCACCGACCAACACGCTTCAGGAACCTTCGCGTCAAGGGGGCG	915
901	CTCTTTCACAGCGGACTTGTGTGGCAGCCATGATTTAGATGGGCAACGTTTGGGGTGCCTACC	960
916	GCTTTCACACAGCGCTTGTGCGCACGCGCATGTCTCAAGATGGGCAACATCTCGCGCTACA	975
961	GGCACCGCGCGGACGATCAGCGCGCAACTGCGCGGGTGCATTAAC	1002
976	GGGACGAGGGCGCAATCAGGCTCAGCTGCTTCAAGGTGAAC	1017

	RESULT 6
AAH44080	ID AAH44080 standard; cDNA: 1158 BP.
XX AC	AAH44080;
XX DT	12-SEP-2001 (first entry)
XX DE	Oryza sativa peroxidase r3025 encoding CDNA SEQ ID NO:19.
XX KM	Oryza sativa; rice; peroxidase; POX; characteristic; gene expression modification; plant; bacterial infection; Magnaporthe grisea; ss.
XX OS	Oryza sativa.
XX FH	Key Location/Qualifiers
FT CDS	14..997
FT FT	/*tag= a
XX PN	/product= "peroxidase r3025"
XX PD	WO200142475-A1.
XX XX	14-JUN-2001.
XX PF	08-DEC-2000; 2000WO-JP08728.
XX PR	10-DEC-1999; 99JP-0352472.
XX PA	(NORO ) JAPAN MIN AGRIC FORESTRY & FISHERIES.
XX OH	Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;

XX	Hiraga S;
PI	WPI: 2001-381695/40.
DR	P-PSDB: AAB99741.
XX	
PT	New set of rice peroxidase genes for analysis of peroxidase expression
PT	in rice under varying conditions and production of rice plants with
PT	desired characteristics
XX	
PS	Claim 1; Page 171-174; 258bp; Japanese.
CC	The present invention describes a set of peroxidase genes found in
CC	plants, especially rice, and their homologues, modified forms and
CC	fragments, where the sequences of the peroxidase genes in the set are
CC	given in AAAH4071 to AAAH4091. Also described are: (1) promoters for the
CC	control of the gene set; (2) the preparation of cassette vectors using the
CC	genes and promoters; (3) analysis of plant characteristics using the
CC	peroxidase set by isolating RNA from the plant, binding the RNA to a
CC	membrane, mixing with a labelled peroxidase gene set, incubating, and
CC	detecting the label signal to show which genes in the set are expressed
CC	in the sample plant; and (4) DNA microarrays for peroxidase gene
CC	expression analysis. The set of genes are used for the analysis of the
CC	pattern of peroxidase gene expression in particular rice plants and
CC	their component tissues and under different environmental conditions,
CC	and modification of rice plants to impart desired specificities of
CC	peroxidase gene expression to impart particular characteristics to the
CC	plants such as response to bacterial infection by Magnaporthe grisea.
SQ	
	Sequence 1158 BP; 255 A; 322 C; 327 G; 254 T; 0 other:
	Query Match            25.0%; Score 338.6; DB 22; Length 1158;
	Best Local Similarity    61.9%; Pred. No. 5,6e-56;
	Matches    580; Conservative    0; Mismatches    339; Indels    18; Gaps
OY	74 CTGCCACCTTGATGCAAGTCCTGGTGCCTCCTCTCTCTGTGTGCGCCACGAC 133
Dd	48 CTGCCGCCCCCTGGTCGARGTGATGTGTGTCTCTCCGCTCCCGCGGGTGTGGCGC 107
OY	134 ACCTCTCGCCCAACGTTCTATGGTGTCTCTCTGCCCCAACCTCGAGACATCGTTGGCGCG 193
Dd	108 AGCTGTGCGCGAGACTTCTAATCTGTACTCTGTGCCCGGAGTGTTCACACGGGTGAACGGGG 167
OY	194 CGATGACCAGCGCGCTGCGCAAGTAGACAGAGATGGGCGCTCTGTGTCAAGGCTTCT 253
Dd	168 GGATGAGTGGCGCATCGCCAGGAGAAGACGCATCGCGCCTTCATCTCCGCTCTCTCT 227
OY	254 TTCACACATCTGCTTCTTCAAGGCTGGAGAGGATCATCTCTCCGAGC----- 301
Dd	228 TTCACACATCTGCTTCTTCAAGGTTGGAGCATCGTGTGCTGTGGAGCACAGGCGAGCT 287
OY	302 CCGGAGGGGGAAGACCGCGCGGCGCCAACCTGAACCTGGTGGCGGCGCTTGTGAJGTGATCG 361
Dd	288 TCACCGGCGGAAGACGGCGAACCCACAACAAGGCTTCGTGACAGAGGTTTGAAGTATCG 347
OY	362 ACAACATCAAGCGGAACGTGAGGCCCGCGTCCCAGGCGTGTGTGTGCGCCGAGATCC 421
Dd	348 ACGCCATCAAGTGGGGGTGAGAACCATCTGTGCCCGCGCTGTCTCTCGCGGAGATCC 407
OY	422 TGCGGCTTCCCGGGCGGCGGAACCAACCTTTCGCGGGCGCGCATGGAGGTCGCGC 481
Dd	408 TTCCCATCTGCTGCGACAGGAACGCTGCCATCTTGGTGTGGGCGGACTGGAGCTGAAG 467
OY	482 TCGGGGCGGGGACTGAGACACGGCCGCTTCTGCGCAACAGCAACCCCGCGCCC 541
Dd	468 TTGGTGGGAGAGACTGCGCGAGCGGCGCTGACGGCGGAACAACACATATCCCGCGC 527
OY	542 CGAGCGCACGCTTGGCAGCGCTATCTCCCTGTGTGGCAAGGCGCTGTGGCGCGCG 601
Dd	528 CGAGCTGCGGAACTGCGCAACCTCATCTCCCTCTTTCGCGCGGCGAGGCGCTTCCAGAGG 587
OY	602 ACAATGAGGCGGTGTGGGGGCGCACAACATCGGGAGAGCGCGGTGACACCTTCCGCG 661
Dd	588 ACAATGGTGCCTCTCGCGATCTTCACACATTITGGGCAAGAGCATGACAAACTTCAGAG 647







(CGR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL), coniferol glucosyl transferase (CGR), coniferin beta-glucosidase (CBG), laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase, caffeic acid methyl transferase, caffeoyl CoA methyl transferase, coumarate CoA ligase, cytochrome P450 1X1A, diaphenol oxidase, flavanone glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase, which are involved in the lignin biosynthetic pathway. The polynucleotides can be used for modulating lignin content, lignin composition and the structure of a plant, especially eucalyptus and pine species, and for modifying the activity of an enzyme involved in lignin biosynthetic pathway, and for producing a plant having altered lignin content, composition, and structure. They can be used for designing probes and primers useful for detecting similar DNA and RNA sequences in any organism and for PCR amplification. The lignin content can be efficiently modified using the polynucleotides. AA67908 to AA68201 and AB16341 to CC AB16449 represent polynucleotide and protein sequences used in the exemplification of the present invention.

Sequence 1224 BP; 343 A; 290 C; 252 G; 339 T; 0 other;

Query Match 20.3%; Score 274.4; DB 21; Length 1224;  
Best Local Similarity 58.1%; Pred. No. 1.2e-43;  
Matches 510; Conservative 0; Mismatches 356; Indels 12; Gaps 1;

QY 128 AGGACAGCTCCGCCAGCTTATGCTGCTCCGCCAGACCTCGACAGCATCCTTC 187  
DB 126 ATGGGAGCTGAGCTCAAGTTTATGCAATGCTGTCCGAGTTGCCATCGATGTA 185  
QY 188 GGGCGGCGATGACCCAGCCCTCGCAAGTGAAGAGATGGGCGCTTCCTGCTCAGGC 247  
DB 186 AATCAGTGTGAGCAACGCGTACGTAAGAGAAAGATGGAGGCTCTTGCGCCG 245  
QY 248 TCTCTTCGACGACTGCTTCGTCAGAGTGCAGGATCGATCCTTCGAGCGC---- 303  
DB 246 TTCACTTTCACATTCCTCTCTCAACGGGTGCCATGCTTATTTAGATGACAGC 305  
QY 304 -----GGAGGGAGAAACACCGCGCGGCACTGAGTGGCGGCTTTGAGG 355  
DB 306 CTACGTTTACCGGAGAAAGACTGACAGCCCAACCTTCGAGAGAGCTTCGAGG 365  
QY 356 TCATGACACATCAAGCGGAGCTCGAGCGCGCTGCCCGCTGCTGTCGCGCG 415  
DB 366 TAATTGACGATTAACAACTCAAGTGAAGCAGCTGACAGTGAAGTGTGTGAG 425  
QY 416 ACATGCTGCGCTTGGCGCGGCGAGCAACCACTTCGCGGCGCGACCTGAGACG 475  
DB 426 ACATTCCTACCTTGTCTGCTGACTCTTGTGAACTTCAAGGCCCAACATGACG 485  
QY 476 TGCCGCTCGGGGCGGAGCTGACAGCGCGCCCTGCTCGCCACAGACACCCG 535  
DB 486 TAATCTTGAAGGAGGAGACTCCAGACTGAGTTTAAGCGCTGCAACACATTC 545  
QY 536 CGCCCGCGAGCGCGCTGCGAGCTCACTCCCTGTTGGGAGGCGGCGCTGCGC 595  
DB 546 CATCTCCCGCTTCCAGCTGAGCACTCATCTATCTTTTCAAGTCAAGGCTTTT 605  
QY 596 CGCGGACATGAGCGCGCTGTGCGCGGACACATCGGGGCGCGGTCACACCT 655  
DB 606 CCAAGACCTTGTGCTCAAGTCTCATCAATTTGGTCAATCATCATCGCTTTT 665  
QY 656 TCCGCGCGCGATCTAGCGGACACGACATCAAGCGCTCTCGGCGGCGTGGCGGC 715  
DB 666 TGAAGACTGAGATCTACAGCAAGAACTTAACCTGCTTTCGTACATCTGTAAG 725  
QY 716 AAGCTGCGCGGCTGCGCGGCGGAGCACTGCGGCCATCGACGTGACAGCGCGG 775  
DB 726 CAAGCTGCGCGGCTGCTGTGCGGAGCAACTCTCTCTTGAATGGGCGACCTTAA 785  
QY 776 TGAGGTGACAGCGGCTACTTCAACACCTGCTGCGGCGGCGGCTGTTCACTGG 835  
DB 786 TCACATTTGACAGCAAGATTAATCTTAATAATACAGAAAGACTTCTCCACTCG 845  
QY 836 ACCAGAGCTCTTCAAGCGCGGCTGCGAGAGCGCGCTGGTGAAGCAGTACAGCGCGC 895

DB 846 ACCAGACCTCTTTAATGAGAGGTTTACAGATTTCTCAGGTTACTGCGTACAGCAATC 905  
QY 896 CCGCTCTTCAAGCGCGCTGCTGAGCGCATGATTTGATGGGCAACGTTGGGGTGC 955  
DB 906 AGACAGCTCTTTATATGACTTACAGCTGCTCATGCTGAGATGGAAATTTAGCCCTC 965  
QY 956 TCACCGGACCGCGCGAGATCAGCGCACTGCGCG 993  
DB 966 TCACGCGACTTACGCGCAATTCGCAAAATCTGCAGG 1003

# RESULT 11

AAH44088 standard; cDNA; 1250 BP.

AAH44088;

12-SEP-2001 (first entry)

Oryza sativa peroxidase c52903 encoding cDNA SEQ ID NO:35.

Oryza sativa; rice; peroxidase; POX; characteristic: gene expression; modification; plant; bacterial infection; Magnaporthe grisea; ss.

Oryza sativa.

Key Location/Qualifiers

CDS 31..1101

/\*tag="a /product="peroxidase c52903"

MO200142475-A1.

14-JUN-2001.

08-DEC-2000; 2000MO-JP08728.

10-DEC-1999; 99JP-0352472.

(NORO) JAPAN MIN AGRIC FORESTRY & FISHERIES.

Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;

Hiraga S;

WPT: 2001-381695/40.

P-PSDB; AAB95749.

Claim 1; Page 220-223; 258pp; Japanese.

The present invention describes a set of peroxidase genes found in plants, especially rice, and their homologues, modified forms and fragments, where the sequences of the peroxidase genes in the set are given in AAH44071 to AAH44091. Also described are: (1) promoters for the control of the gene set; (2) the preparation of cassette vectors using the genes and promoters; (3) analysis of plant characteristics using the peroxidase set by isolating RNA from the plant, binding the RNA to a membrane, mixing with a labelled peroxidase gene set, incubating, and detecting the label signal to show which genes in the set are expressed in the sample plant; and (4) DNA microarrays for peroxidase gene expression analysis. The set of genes are used for the analysis of the pattern of peroxidase gene expression in particular rice plants and their component tissues and under different environmental conditions, and modification of rice plants to provide desired specificities of peroxidase gene expression to impart particular characteristics to the plants such as response to bacterial infection by Magnaporthe grisea.

Sequence 1250 BP; 227 A; 433 C; 375 G; 215 T; 0 other;



Query Match	20.28;	Score 273;	DB 22;	Length 1250;
Best Local Similarity	58.48;	Pred. No. 2.2e-43;		
Matches 530; Conservative	0;	Mismatches 350;	Indels 27;	Gaps 2

QY	126	CCACGACACAGTCCTGCGCCACAGCTTTCATACGCTGCTCCGACCCCAACCTGAGAGCAATGCT	185
Db	105	CCGCGGCGACAGTGTGAGGTGGGGGTTCTACAAACACACAGCTGCCCAATGCCAGAGAGCTGCT	164
QY	186	TGGGGGCGGCATGACACCAAGGCGCTGCCAAGTGTAGCAGAGAGATGGGGCGCTCTGTGCTCAG	245
Db	165	CCGGAGAGGCGGTACACAAAGGCTTTCGGCACAGCACTCCGGATATCCGCGCCGCTCATCCG	224
QY	246	GCTCTTCTCCACACATGCTCTTGTTTAAAGGCTGCACAGCGAATCGATCTCTT-----	296
Db	225	CCCTCATTTTCCAGACATGCTTCTTGTCAGAGGTTTGGCAGCGCTGGTGGCTGTGACGTCGCC	284
QY	297	CCAGCGCCGGAGGGAGAAAGACCGCCCGGGGCCAACTGAATCTGGGTGGCGGCGCTTTAGGT	356
Db	285	CACACAAACAGCGGGAGAGCGCAGCGGCGCGCAAGAACCCACCTCCGTGGTGGCTTCCAGGT	344
QY	357	CATCGACACATCATCAAGGGAGACGTTCAGAGCCGCGCTGCCCGCGCTCTGTGCGCCCGCA	416
Db	345	GATCGACGCCCGCCCAAGGCGCCCGCTGCAGAGAGGTGCGCGCGACGGGTGCTCGCCGCGA	404
QY	417	CATTCCTCGCGGTCGCGGGGCGAGCGAAGCAACTTTCGGCGGGCGCCACCTGGAAGGT	476
Db	405	CATTCGTCGCGTTCGCGCGCGCGCACAGCTCAACTCAACGGCGGCGTCTCTCATCAGGT	464
QY	477	GCGCGTCGGGCGGAGACTCGACGACGAGCGCCACGCGCTCGTTCGCCAACAGAACCCGCC	536
Db	465	CCCTCTCGGCGCGCGCACAGGCAACGTCTCCGTCCCGACGACCGCATCGACACACCTCC	524
QY	537	GCGCCCGACGCGCACAGCCCGGACAGGTCATCTCCGTTGGCGGAGGAGGGCGCTGCGCG	596
Db	525	CCAGGCCACACTTCAACCGCGCCGACAGTCTGCGCACAGTCTGCCAACAAAGTCCGTACCGC	584
QY	597	GCGCGCATGACAGGCGCTGTTCGGGCGCGCACACATCGGGAGGGCCCGGTGCACACGTT	656
Db	585	CGAGGAGATGGTCTCTCTTCGCGGCGCCACACGCTCGGGCGCTCTTGTGCTCTCTT	644
QY	657	CCGCGGCGCGCATCTACGCGCACACACATCAATCAACGCTCTC-----TT	698
Db	645	CCTCGCGCGCATCTGGAACAAACACACCCCATCTGTGAGACAGGGGCTGAGCCCGGGGTA	704
QY	699	CGCGGCGCTGGGGGAGAGACAGAGCGCCCGGGGTTCGGCGGAGCAGGCACTGGCGCCCAT	758
Db	705	CGGGGCGCTGTGAGGGGCGTGTGCGCGCTGAGACCGGTGCGGAGCGAGCGAGCGAGT	764
QY	759	CGAGGTGAGACCGCGGTGAGGTTTCACACGCGCTACTTCAACAACTCTGCTGCGGCGC	818
Db	765	CGACGTGAGACCGCGCGCGAGCGCTGTGAGAACAACTACTACAGGTGTGTCGCGCTCAACT	824
QY	819	GCGCGCTTTCACACTCGGACCGAGAGGCTCTTCAACGGCGGGGTGCGAGGACCGGTGGTAG	878
Db	825	GGGGGCTCTTCTTCTTCGACCAACAGCTGGGGGTGAACCGACGCTGGGCGCGTGGGTGAG	884
QY	879	GCAGTACAGCGCCAGCGCGCTCTGCTCTTCAAGCGCGACTCTGTGGCACCACTGATTAGAT	938
Db	885	CAGCTTTCGCGCGCAACAGAGAGCGCTGTGGAAGAGGAAGTCTCTCCGCGCATGCTCAAGAT	944
QY	939	GAGCAACCTTGGGGGTGCTACCGGCAACCGCCGCGAGATGATGAGGAGCAACTGCCGGGTCT	998
Db	945	GGGGAGCATGAGGTGTCTCACCGCGACCGAGGGCAGGGTCTAAGCTGACAGCTGCT	1004
QY	999	CACACGC	1005
Db	1005	CACACAC	1011

RESULT 12  
AAA68192  
ID AAA68192 standard; DNA; 801 BP  
XX

	AA68192.	
Df	24-OCT-2000	(first entry)
Xx		
Pt	<i>Pinus radiata</i> peroxidase nucleotide sequence SEQ ID NO:368.	
Rw		
Km	Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;	
Rw	<i>Pinus radiata</i> ; Monterey pine; ds.	
Xx		
Oz	<i>Pinus radiata</i> .	
Fn	WQ20002209-A1.	
Pd	20-Apr-2000.	
Xx		
Pf	06-Oct-1999; 99WO-NZ00168.	
Xx		
Pr	09-Oct-1998; 98OS-0169789.	
Pr	14-Jul-1999; 99US-0143811.	
Pa	(GENE-) GENESIS RES & DEV CORP LTD. FLETCHER CHALLENGE FORESTS LTD.	
Xx		
Pt	Bloksberg LM, Hayakkala JU:	
Dk	WPJ: 2000-317962/27.	
Pt		
Pt	Novel polynucleotide encoding enzymes involved in lignin-biosynthetic pathway useful for producing transgenic plants especially eucalyptus and pine species having altered lignin content, composition and structure -	
Pt		
Ps	Claim 1; Page 189; 213pp; English.	
Cc	The present invention describes isolated polynucleotides and proteins encoding and representing the enzymes cinnamate 4-hydroxylase (C4H), coumarate 3-hydroxylase (C3H), phenolase (PNL), O-methyl transferase (OMT), cinamoyl alcohol dehydrogenase (CAD), cinamoyl-CoA reductase (CCR), phenyllalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (CL), coniferol glucosyl transferase (CGT), coniferyl beta-glucosidase (CBG), laccase, peroxidase, ferulate-5'-hydroxylase (FSH), alpha-amylase, caffeic acid methyl transferase, caffeoyl CoA methyl transferase, coumarate CoA ligase, cytochrome P450 1XII, diphenol oxidase, flavanol glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase, which are involved in the lignin biosynthetic pathway. The polynucleotides can be used for modulating lignin content. Lignin composition and the structure of a plant, especially eucalyptus and pine species, and for modifying the activity of an enzyme involved in lignin biosynthetic pathway, and for producing a plant having altered lignin content, composition and structure. They can be used for designing probes and primers useful for detecting similar DNA and RNA sequences in any organism and for PCR amplification. The lignin content can be efficiently modified using the polynucleotides. AA67908 to AA68201 and AB16341 to AB16449 represent polynucleotide and protein sequences used in the embodimentation of the present invention.	
Sq	Sequence 801 BP; 195 A; 217 C; 195 G; 194 T; 0 other;	
	Query Match	19.8%; Score 268.6; DB 21; Length 801; Best Local Similarity 60.7%; Pred. No. 1.5e-42; Matches 464; Conservative 0; Mismatches 289; Indels 12; Gaps 1;
Gy	93 CTGTGGCCGCAGTTTCCTCTCTCTTGATGCCACCAACAAGACTTCCGACCATTCTTA	152
Dd		
	36 CTTATATCGCTTGATGTGTTGCAGCGTTGCAGACGCGAGCATGTTCACCAACGTTTTA	95
Gy	153 TGCGTCTCTCGCCCCAACCTGCAAGAAGATGGTTCGGCGGCAGTAGCACCAAGCGCGTGC	212
Dd		
	96 TAATACTATCTGTCCCAAACACTCTCAAGCGTGAAGCGCCGCTGTGAAGCAACGGTGGC	155
Gy	213 AAGTAGACAGAGATGGCGCCTCTCTCTCAGAGCTCTTCTTCACAGATGCTTGGTTCA	272
Dd		
	156 CAACGAAACACCGATGGGGCATTCCTCTCCGCGCTTCACACTTCACAGACGCTTGGTTAA	215

Qy	273	AGCGTCGACGGATGATGATCTCTTTCGACGGC-----GGAGGGGAGGAAGACCG	320
Db	216	TGCTTCCGATGGGCTGTGTGCTGTGGACGACTCTTGACTCTTAAGTGGAGGAAGACCG	275
Qy	321	CGGGCCGAACCTGAACTCGTGCGCGGGCTTGTAGGTCAATGACACCATCAAGCGAACT	380
Db	276	CGTTCCCAACCAACAACTTGGCGCAAGGGGTTCGATGTGAATGACACATCAACATCTCAAGT	335
Qy	381	CGAGGCGGCGTCCCGGCGCGTGTGTGTGTCGCGCATCTCGGCTGTGCGCGCGCA	440
Db	336	GGAGAGAGTTTCACTAGTGGAGTGTGTGTGTGTCGCGAGATATTTGGCTATTGCTGCTAAGA	395
Qy	441	CGGACCAACCTTCTCGCGCGGGCCGACTGTGAGCTGTCCGTGCGGCGCGGACTCGAC	500
Db	396	TTCCTGTGTTCACAGTTGGGAGGGCCCAACATGGAAGTGAAGCTGTGGGAGAGACTCGAC	455
Qy	501	GACGGCGACGGCTGCTGTGCGCAACAGACACCCCGCCCGGAGGGCGACACTCGGCAC	560
Db	456	GACTGCGACACCTAAGTGGTGTGCAAAACACAACTTCGGGCTCTTACTATTCTCAATGC	515
Qy	561	GCTCATCTCCCTGTTCCGCGAGGAGGCGCTGTGCGCGCGACATGACGCGCTGTGGG	620
Db	516	TCGCATCTCATATTATTTCAAGCTCAAGGCTCTTTCACAGGAAGACATGTTGTCTCATAGG	575
Qy	621	CGCGCACACATCGGCGAGGCGCGGTGCACACACTTTCGCGGCGCGATCTACAGGCGACAC	680
Db	576	TGGCGACACCATAGGCGCAAGCGCGGTGCACAAAGCTTCAGAGGCGCGCATCTACAGCAATC	635
Qy	681	CGACATCAACGCGCTCTTTCGGGGGCGCTGCGGCGACAGACGTGCGCGGCTCCGGCGCGA	740
Db	636	CACATCTTAATGAGATACGCAACTTCTCCGTGAAGACAAACTGTGCACTACAGGAACGGA	695
Qy	741	CGGCAACCTGGGCGCCATCGACTGCGAGAGCGCGGTGAAGTTCGACACGCGCTACTTCAC	800
Db	696	CACACACCTGTACCATTTGGATGCTGTACTTCCCACTACGTTTGACATCACTACTACTC	755
Qy	801	CACACTCTGTGTGGCGGGGGCGTGTGCATCGAGGACAGAGAGCT	845
Db	756	AAATCTGACAGCCAAAGGAGACTTCTCCACTCCGACGACGACGCT	800

```

RESULT 13
AAA68176
ID AAA68176 standard; DNA; 1391 BP.
XX
AC AAA68176;
XX
DT 24-OCT-2000 (first entry)
XX
DE Eucalyptus grandis peroxidase nucleotide sequence SFO ID NO:352
XX
KW Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;
KM Pinus radiata; Monterey pine; ds.
XX
OS Eucalyptus grandis.
XX
PN MO20002209-A1.
XX
PD 20-APR-2000.
XX
PF 06-OCT-1999; 99MO-NZ00168.
XX
PR 09-OCT-1998; 98US-0169789.
XX
PR 14-JUL-1999; 99US-0143811.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN, Havukkala IJ;
XX
WP1: 2000-317962/27.
XX

```

Novel polynucleotide encoding enzymes involved in lignin-biosynthesis  
 pathway useful for producing transgenic plants especially eucalyptus  
 and pine species having altered lignin content, composition and  
 structure -

The present invention describes isolated polynucleotides and proteins encoding and representing the enzymes cinamamate 4-hydroxylase (CH), cinamate 3-hydroxylase (CH3), phenolase (PNL), O-methyl transferase (OMT), cinemayl alcohol dehydrogenase (CAD), cinamoyl-CoA reductase (CCR), phenylalanine ammonia-lyase (PAL), 4-cinnamate:CoA lyase (4CL), coniferyl glucosyl transferase (CGT), coniferin beta-glucosidase (CBG), laccase peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase, caffeic acid methyl transferase, caffeoyl CoA methyl transferase, cinamoyl CoA lyase, cytochrome P450 1XX1, diphenol oxidase, flavanol glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase, which are involved in the lignin biosynthetic pathway. The polynucleotides can be used for modulating lignin content, lignin composition and the structure of a plant, especially eucalyptus and pine species, and for modifying the activity of an enzyme involved in lignin biosynthetic pathway, and for producing a plant having altered lignin content, composition and structure. They can be used for designing probes and primers useful for detecting similar DNA and RNA sequences in any organism and for PCR amplification. The lignin content can be efficiently modified using the polynucleotides. AAs67908 to AAs68201 and AAs16341 to AAs16449 represent polynucleotide and protein sequences used in the exemplification of the present invention.

50 Sequence 1391 BP; 399 A; 321 C; 307 G; 364 T; 0 other;

Query Match	19.1%;	Score 258.6;	DB 21;	Length 1391;
Best Local Similarity	56.7%;	Pred. No. 1.3e-40;		
Matches 505; Conservative	0;	Mismatches 374;	Indels 12;	Gaps 1;

QY	124	GGCCAGCAGAGCTTCGCGCCCGGTCGATGACGCTCCGCGCCCAACCTGAGAGCATC	1830
	125		
Db	300	GCCCGTGTAGAGCTCACACCGAGTCATTATCAACAATGTGCCAAGAAAGCATTTGTCATT	3590
QY	184	GTTGCGGCGGCATGACCCAGGCCCTCGCAAGTAGACAGAGATGGGGCCTCTCTGTC	2430
	185		
Db	360	GTTGACACTGAGATACAAAACCAATCAAGAATGAGACCCGGAGCGGCTCTTGCTT	4190
QY	244	AGGCTCTTCTTCACAGCTGCTTCTGTTCAAGGCTGCGAGCGATCGATCCTTCTCGA----	2990
	245		
Db	420	CGGCTGCATCTCATGCTGCTTCCTCATATGGGTCATGCGCTGATATTGTTGGATTGAC	4790
QY	300	-----GGCGGAGGGGAGAAAGCGCGCGGCGCAACTGCACTGAGTGGGGGGCTT	3510
	301		
Db	480	ACGCTTAGCTTGTGGGGGAGAAAACAGACACTGCCAACACAAATTCCTGTAGAGGTTG	5390
QY	352	GAGTCATGACACCATCAAGCGGAAAGCTGAGAGCGCGCTGCCCGGCGCTGTGTGTC	4110
	353		
Db	540	GAATGATGACGCGCATCAAGGCTTAGTGTGAAAGAGTGGCCGTGGATGGTTTTCCTGT	5990
QY	412	GCCGACATCTCGCGCTTCCGCGCGGCGAGCAGCAACCTTCTGCGGCGGCTGACCTGG	4710
	413		
Db	600	GCAATATGCTGGCCCTGGCTGCTCGGACTCAAGTGTTCATTATTTGGAGAGTCTTCATGG	6590
QY	472	AGGCTGCGCGCTGGGGCGCGGAGACTCGACAGACGGCAGCGCTCTGCTGCCAACAAGCAAC	5310
	473		
Db	660	ACGGTAAGCTTAGGGGAAGAAAGGATTTCCATTATCTGTAGCAGGAGCCTTGCTTAACCTCC	7190
QY	532	CCCCCGCCCCGAGGCCAGCCTGCGGACAGCTCATCTCCCTGTTGCGGAGGCGAGGGCTG	5910
	533		
Db	720	ATACTTCACACTACTTATATCTAGTAGTCTCTATAACAGAGCTGGCTGCTACAGGTCCTT	7790
QY	592	TTCGCGGCGGACATGACGGCGGTGTCGGGGCGGACACCATCGGGGAGGGCCGGGTGAC	6510
	593		
Db	780	TCACTCAGAAACATGTGTGGCTTTTCTGTGTTACATATCCATTGGCTTAGCGAGATGCACT	8390
QY	652	ACCTTCGGGGCGGACTAGCGGAGACACGACATCAAGCCTCTCTTGGCGGCTCGCG	7110
	653		



Db 840 TCCTCCGAAGAGGATCTACACGACTCGACATGATACATCTCTGCCCATTAATTG 899  
 QY 712 CAGCAGACGTGCCCGGCTCGGCGGCGGACGCAACCTGGCCCATCGAGTCGAGACG 771  
 Db 900 CAGAAAGTATGTGCCAGGATGGAATGATGATGTCTTCAAGGCTAGACATCCAAAGC 959  
 QY 772 CCGGTGAGTTCGACAGCGGCTACTTCACACGCTGTCGCGGCGGCGGCTGTTCAC 831  
 Db 960 CCGACCTCTTGTACACACCTTACTACACATTTACTGCAAGAAAGGCGCTTCTTAC 1019  
 QY 832 TCGACACGAGAGCTCTTCAAGCGGCGGTCGACGAGCGCGCTGTGAGCAGTACGCGC 891  
 Db 1020 TCTGATCAAGAGCTCTTCAAGGAGTCTGTGATTCACCTGTCGCAAGAAAGTATGATGC 1079  
 QY 892 AGCGCTCTGCTCTTCAAGCGGCTGCTGTCGACGATGATTAAGATGGGCAACGTTGG 951  
 Db 1080 GACACAGGAAATTTTTCGAGATTTTGCCAGGCAATGATCAAAATGAGCAATTAAG 1139  
 QY 952 GTGTGACGCGGACGCGGACGATGACGCGGCACTGCGGCGGTGAC 1002  
 Db 1140 CCCCCAAGAGGAGCATGCTCAATTAAGAAATTTGACGAAAGTGAAC 1190  
 RESULT 14  
 AAF62772 standard; DNA; 975 BP.  
 AAF62772:  
 03-MAY-2001 (first entry)  
 Synthetic Nicotiana sylvestris peroxidase gene.  
 XX Tobacco: Nicotiana sylvestris; peroxidase; insecticide;  
 KW transgenic plant; insect-resistant plant; Ostrinia nubilalis;  
 KW Heliothis zea; Spodoptera frugiperda; ds.  
 XX Nicotiana sylvestris.  
 OS Synthetic.  
 PN WO200109353-A1.  
 PD 08-FEB-2001.  
 PF 28-JUL-2000; 2000WO-EP07310.  
 PR 30-JUL-1999; 99US-0365150.  
 PR 16-NOV-1999; 99US-0165683.  
 PA (SYNG-) SYNGENTA PARTICIPATIONS AG.  
 PI Lagrimini IM, Desai NM;  
 DR WPI: 2001-182962/18.  
 PT Novel anionic peroxidase genes isolated from Nicotiana tomentisiformis  
 PT and anionic peroxidase protein encoded by the genes, useful for  
 PT producing plants resistant to insects -  
 PS Claim 19; Page 54; 58pp; English.  
 CC The present sequence is a synthetic Nicotiana sylvestris peroxidase  
 CC gene which is optimised for expression in plants. The sequence may be  
 CC introduced into plants to produce insect-resistant plants, preferably  
 CC maize. It is useful for controlling insects such as Ostrinia nubilalis,  
 CC Heliothis zea and Spodoptera frugiperda. The transgenic seeds and plants  
 CC can be used for the breeding of improved plant lines which, for example,  
 CC increase the effectiveness of conventional methods such as herbicide or  
 CC pesticide treatment or allow the conventional methods to be dispensed  
 CC with. The peroxidase enzymes can be used in multiple insect control  
 CC strategies, resulting in maximal efficiency with minimal impact on the  
 CC environment.  
 CC

SQ Sequence 975 BP; 198 A; 373 C; 267 G; 137 T; 0 other:  
 Query Match 18.8%; Score 254.4; DB 22; Length 975;  
 Best Local Similarity 58.3%; Pred. No. 8,1e-40;  
 Matches 541; Conservative 0; Mismatches 351; Indels 36; Gaps 4;  
 111 CCTCTGCTGTGCGCCGACGACGCTCTGCGCAGCTCTGATGAGCTCTGCGCCCA 170  
 Db 45 CATCTTGCGCGCCAGACACGCGCGAGCTGAGCGCACTTCTACGACACGCTGCCCA 104  
 QY 171 CCGCAGACATCTGTCGCGCGCGGATGACCCAGCGCGCTGCGAAGTGAAGAGAGG 230  
 Db 105 CGGACACGATCGTGTGCGCGCGGATGAGACGCGCGCGCGCGCGCGCGCG 164  
 QY 231 CGCCTCTGCTGACGAGCGCTCTGCAAGCTGCTCTGTTCAAGGCTGCGACGAGAT 290  
 Db 165 CGCAGATCATCCGCTGCACTTCCAGCTCTTGTGAACGCGCTGCGAGCGACAT 224  
 QY 291 CTTTTCGAGCGCGGAGGAGGAGAGGAC-----GCGGCGGAGAACGTAAGTGGG 344  
 Db 225 CTTGCGACACCGAGCGACCGACCGAGCGAGCGAGCGCTCTGCAAGCTGGGCGG 284  
 QY 345 CGGCTTGAAGTCAATGACACATCAAGCGAAGCTGAGGCGCGCTGCCCGCTGCT 404  
 Db 285 CGGCTTGAAGTCAATGAGAGACATCAAGAGCGCGCTGAGAAAGTGGCGAGTGG 344  
 QY 405 GTGTCGCGCGACATCTGCGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 464  
 Db 345 GACCTGCGCGACATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 404  
 QY 465 GACCTGAGCG 524  
 Db 405 TACCTGCGAGTGTCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 464  
 QY 525 CAGCAACCG 584  
 Db 465 CAGCGACATCG 524  
 QY 585 GGGCTGTGCG 644  
 Db 525 GGGCAATGAGATCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 584  
 QY 645 GTGACACACTTCG-----CGGCGCGATCTAGCGCGACCGA 683  
 Db 585 CTGCGGCACTTTCGAGGAGAGATGTTCAACTTCAAGCGCGCGCGCGCGCGCGCG 644  
 QY 684 CATCAAGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 743  
 Db 645 CGTGAAGCGCACTTCTGCGAGACCTTCAGGGCATCTGCGCGCGCGCGCGCGCG 704  
 QY 744 CAACGTGGCG 800  
 Db 705 CAACACCTTTCACCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 764  
 QY 801 CAACGTGCTGTGCG 854  
 Db 765 CAACCTTCAGAGCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 824  
 QY 855 CGGATGCGAGAGCG 914  
 Db 825 CAGCGCACCATCG 884  
 QY 915 CTTGTCGAGCG 974  
 Db 885 CTTGTCGAGCG 944  
 QY 975 GATCAGCG 1002  
 Db 945 GATCAGCG 972  
 RESULT 15  
 AAF62773



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 08:13:26 ; Search time 50 Seconds  
(without alignments)  
8304.807 Million cell updates/sec

Title: US-10-047-825-3

Perfect score: 1354  
Sequence: 1 aattcgccagcagcttaagc.....aaaaaaaaaaaaaaaa 1354

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : Issued Patents.NA:\*

1: /cgn2\_6/prodata/1/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/lna/5A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/lna/5B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/lna/PCRTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	287.2	21.2	1171	4	US-09-615-192A-369 Sequence 369, App
2	274.4	20.3	1224	4	US-09-615-192A-363 Sequence 363, App
3	268.6	19.8	801	4	US-09-615-192A-368 Sequence 368, App
4	258.6	19.1	1391	4	US-09-615-192A-352 Sequence 352, App
5	254.4	18.8	975	4	US-09-365-150-3 Sequence 3, Appl
6	245	18.1	975	4	US-09-365-150-4 Sequence 4, Appl
7	234	17.3	916	4	US-09-615-192A-361 Sequence 361, App
8	226.6	16.7	1354	4	US-09-097-319A-5 Sequence 5, Appl
9	182.8	13.5	586	4	US-09-615-192A-362 Sequence 362, App
10	175.8	13.0	646	4	US-09-615-192A-365 Sequence 365, App
11	121.6	9.0	659	4	US-09-615-192A-359 Sequence 359, App
12	119.6	8.8	1072	1	US-08-190-028A-9 Sequence 9, Appl
13	119.6	8.8	1072	2	US-08-462-695-9 Sequence 9, Appl
14	119.2	8.8	1522	4	US-09-615-192A-371 Sequence 371, App
15	113	8.3	1073	4	US-09-615-192A-370 Sequence 370, App
16	110.8	8.2	461	4	US-09-615-192A-375 Sequence 375, App
17	102.8	7.6	364	4	US-09-615-192A-366 Sequence 366, App
18	102	7.5	474	4	US-09-615-192A-373 Sequence 373, App
19	101	7.5	519	4	US-09-615-192A-364 Sequence 364, App
20	95	7.0	278	4	US-09-615-192A-89 Sequence 89, Appl
21	93.8	6.9	1326	2	US-08-671-320-12 Sequence 12, Appl
22	93.8	6.9	1326	2	US-08-868-577-12 Sequence 12, Appl
23	93.2	6.9	955	4	US-09-615-192A-355 Sequence 355, App
24	91.8	6.8	278	2	US-08-713-000-13 Sequence 13, Appl
25	91.8	6.8	278	2	US-08-975-316-13 Sequence 13, Appl
26	91.8	6.8	278	4	US-09-211-710-13 Sequence 13, Appl
27	91.8	6.8	278	4	US-09-615-192A-13 Sequence 13, Appl

28	89	6.6	1314	2	US-08-868-577-10 Sequence 10, Appl
29	89	6.6	1315	2	US-08-671-320-10 Sequence 10, Appl
30	83.2	6.1	717	4	US-09-615-192A-350 Sequence 350, App
31	83	6.1	6550	4	US-09-097-319A-1 Sequence 1, Appl
32	81.6	6.0	1131	1	US-08-181-271A-22 Sequence 22, Appl
33	81.6	6.0	1131	1	US-08-449-315-22 Sequence 22, Appl
34	81.6	6.0	1131	1	US-08-444-803-22 Sequence 22, Appl
35	81.6	6.0	1131	1	US-08-449-043-22 Sequence 22, Appl
36	81.6	6.0	1131	1	US-08-456-265A-22 Sequence 22, Appl
37	81.6	6.0	1131	1	US-08-455-416-22 Sequence 22, Appl
38	81.6	6.0	1131	1	US-08-455-244-22 Sequence 22, Appl
39	81.6	6.0	1131	1	US-08-454-876-22 Sequence 22, Appl
40	81.6	6.0	1131	2	US-08-457-364-22 Sequence 22, Appl
41	81.6	6.0	1131	2	US-08-456-262-22 Sequence 22, Appl
42	81.6	6.0	1131	2	US-08-455-240-22 Sequence 22, Appl
43	81.6	6.0	1131	2	US-08-455-736-22 Sequence 22, Appl
44	81.6	6.0	1131	2	US-08-971-217-22 Sequence 22, Appl
45	81.6	6.0	1131	4	US-09-350-600-22 Sequence 22, Appl

#### ALIGNMENTS

RESULT 1  
US-09-615-192A-369

/ Sequence 369, Application US/09615192A  
/ Patent No. 6410718

GENERAL INFORMATION:  
/ APPLICANT: Blosberg, Leonard N.  
/ APPLICANT: Havukala, Ilka

/ TITLE OF INVENTION: Materials and Methods for the  
/ FILE REFERENCE: 11000.1003c4U

/ CURRENT FILING DATE: 2000-07-12  
/ PRIOR FILING DATE: 1997-11-21

/ PRIOR APPLICATION NUMBER: US 08/713,000  
/ PRIOR FILING DATE: 1996-09-11

/ PRIOR APPLICATION NUMBER: US 09/169,789  
/ PRIOR FILING DATE: 1998-10-09

/ NUMBER OF SEQ ID NOS: 405  
/ SOFTWARE: FastSeq for Windows Version 3.0

/ SEQ ID NO 369  
/ LENGTH: 1171

/ TYPE: DNA  
/ ORGANISM: Pinus radiata

US-09-615-192A-369

Query Match 21.2%; Score 287.2; DB 4; Length 1171;  
Best Local Similarity 59.7%; Pred. No. 2.4e-51;  
Matches 528; Conservative 0; Mismatches 338; Indels 18; Gaps 2;

QY	128	ACGCACACCTCCGCCAGCTTCTATGCGTCTCTGCCCACTCGACGATCGTTC	187
DB	119	ATGGCAGCTAGTCCGCCAGCTTCTATGCGATCTTGGCCAGCAGCTGCGTAGTGA	178
QY	188	GGGCGGCATGACCCAGCGCTCGAAGTGAAGAGATGGGCGCTCTGCTGAGGC	247
DB	179	AGCGCGAGTGAAGCAGCGTCTGCTAAGCAAGAAAGATGGGCGCTTGTGCGCC	238
QY	248	TCCTTTCACGACTGCTTCTGTAAGGCTCGAGGATCGATCTTCTCGAGC-----	301
DB	239	TGCACTTTCACGACTGCTTCTGTAAGGCTCGATGAGGCTCGTCTGAGCAATCTT	298
QY	302	-----CCGAGGGGAGAGAGCCCGCGGAGCAGTCTGCTGCTTTAGG	355
DB	299	CGACCAATTAAGTGGGAGAGACACTTAATCCCAATTCGCGAGGATTCAGC	358
QY	356	TCATCGACACCATCAAGCGAGAGCGTGGAGCGCGCGCGGCGTGTGCGGCGC	415
DB	359	TAAATAGATACCATTAAGAGCAATGTCAGAAAGCTTGCAGTAGTGTCTCTCTG	418

QY	416	ACATCTCTCGGCTTGCCCGCGCGGACGAGAACCAACTCTCTCGGCGGGCGAGCTGAGGG	475
Db	419	ACATTCTCGCATTTGCTGCTCGTGTATTTCTGTGTAACTGGGGCGGTCTCTTATGGACAG	478
QY	476	TGGCGCTTCGGGCGGCGGAGACTTCGACGACGGCCAGCGCTCTGCTCGGCACAGCAACCCCC	535
Db	479	TAAATTGGGAAGGCGGAGACTTCGACACACACTTACAAAAGCGGTGCGAACAAGTAAATTC	538
QY	536	CGCCCCGAGACGGCCAGGCTCGGCGACGCTATCTCCCTGTTGGCAGAGCAAGGCGCTGTCC	595
Db	539	CGCCTCGAGACTTCAGTCTGAGCAACCTCATCTACATATTCACAAGCGAGGAGACTCTCCG	598
QY	596	CGCGCGACATGACGGCGCTGTGCGGGCGCGACACACATCGGGCAGGGCCCGTGACACACT	655
Db	599	CAAAAGAAATGGTTGCACTTTCTGGCGGTATACCATTCGGGCAGGGCGCAATGCAAGAATT	658
QY	656	TCGCGGGCGGATCTACGGGCGACACGCACATCAACGGCTCTTCGGGCGCTCGCGGACG	715
Db	659	TCAGAGCCCATATTTTACACGACGACCAACATAGACAGTGGTAGGCCACTTANTTGGTT	718
QY	716	AGAGCTGCCGCGGCTC-----CGGCGGCGAGGCGCAACTGGCGGCCCATCGACGTGCAGA	769
Db	719	CAAAAGTGCAGAGTACACACAGCGCTCGGAGACAGCACTGTGCGCAATTGATTATATGTA	778
QY	770	CGCGGCTAGGTTGGACACGGGCTACTTTCACCAACCTGCTGTGCGGGCGGGCGCTGTCC	829
Db	779	CTCCCACTGTGTTTGAACAAAACATTTTACACGCGACGCAAGAAAGCCAAAAGAACTCTCC	838
QY	830	ACTCGGACACAGAGCTCTTCAACGGCGGGTGCAGGACGGCTGGTGTAGGACAGTACAGCG	889
Db	839	ACTCGGACACAGAACTCTTCAACGCGAGGCTCCACTGATTTCAAGGTACATACGTAGCGCT	898
QY	890	CGAGCGCTCTGCTTTCAACGCGCACTTCTGTGACGCACTGATTTAGGATGGCGACAGTTG	949
Db	899	CCAAACAGAAACACTTCTTCTCCGATTTTGTGCGGCACTGTTAAAGATGGGAAATATCA	958
QY	950	GGGTGCTCAACCGGACCGCGCGAGAGATCAAGGCACTGCGG	993
Db	959	AACCTCTTACGGGACACGCGGACAGATCTCCAAAGAACTGAGG	1002

RESULT 2  
US-09-615-192A-363  
: Sequence 363, Application US/09615192A  
: Patent No. 6410718

```

; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.

```

```

; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; PUBLICATION NO.: 1,000, 1,003, 47

```

FILE REFERENCE: 11000.1003C4U  
CURRENT APPLICATION NUMBER: US/09/615,192A  
CURRENT FILING DATE: 2000-07-12

;; CURRENT FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 08/975,316  
;; PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11

;; PRIOR APPLICATION NUMBER: US 09/169,789  
;; PRIOR FILING DATE: 1998-10-09

```

; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 363
; LENGTH: 1224
;

```

```

; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-363

```

Query Match	20.38%	Score 274.4:	DB 4:	Length 12242
08 09 010 1224 203				

Best Local Similarity	58.1%;	Pred. No. 1.1e-48;
Matches 510;	Conservative	0;
Mismatches	356;	Indels 12;
Gaps	1;	

128 ACCGACAGCTCTGCCCCAGGTTCTATGCGTCTCTGCCCCAACCTGCAGAGCATCGTTC 187

D	126	ATGGCAGCTGAGCTCAACGTTTTATGCCAAATGCTGCCAGGTGGCCATGATAGTGA	185
Q	188	GGGGGGGATGACCCAGGCGGTGCAGAGTACAGAGAGATGGCGCTCTCTCTCTCAGGC	247
D	186	AATCAGTGGTGAACAGCGGTACTTAAGAGAAAMAAAGGAGGGCTCTTGTCTCGCC	245
Q	248	TCCTCTCCAGACTGCGTGGTTCGAAGCGTCGACGGATGATCTCTCTGACGCG----	303
D	246	TTCACTTTCACGATTGGCTCTCAGCGGGTCCGATGGTTCAATCTTATTGTGATGACAG	305
Q	304	-----GGAGGGAGAAAGACCGCGGGCCGAACCTGAACTCGTGGCGGGCTTTGAGG	355
D	306	CTAGGTTACGGGAGAAAGACTCAGAGCCCAAGCGCAATTCTGCGACGAGGCTTGAGG	355
Q	336	TCATGACACCATCAAGCGGAGCTCGAGGCGCGTCCCGCGGGGTGTGCTGGCGCG	415
D	336	TAAATGAGAGATTAAACCTCAAGTGGAGGAGGCTCAGTGGAGTGCTGTCTGTGACG	425
Q	416	ACATCTCGGCGCTTGCCGCGCGGACAGGAACCAACTTCTCGGCGGGCGCAGCTGGAGC	475
D	426	ACATTTCTACCATTTCTGCTGTGACTCTATTTGTAACCTCAAGGCCAACAATGGACG	485
Q	476	TGCCGCTCGGGCGGGGACTTGCAGACGGCCAGCGCTGCTGTGCCAAGACACCCC	535
D	486	TAAATCTTGAAGGGGAGACCTCCAGACTGCGAGTTTAAAGCGCTGCAAAACCAACTTC	545
Q	536	CGCCCCGACGGCCAGGCTTGGGACGCTATCTCCGTTTGGCAGGACAGGCTGTGCG	595
D	546	CATCTCCCGCTTCCAGTGTGACACACTCATTCTATCTTTCAAGCTCAGCGTCTTTCTA	605
Q	596	CGCGGACATGACGCGCTGTGCGGCGGCGACACATCTGGGACGCGCGGCTGACCACT	655
D	606	CCAAAGACTTGTTCACCTCAGGTGTCTATTAATGTGTCATTCAGATGCGCCCTTTT	665
Q	656	TCCGGCGCCGATCTACGGCGACACCGCATCAACGCTCTCTTGGCGGCGGTGGCGAGC	715
D	666	TCAGAACTCGGATCTACACAGAAAGAACTTAAAGCTGCTTGGCTACATCTTAAAG	725
Q	716	AGACGTGCGCGGGGTCGCGGCGGCGACGCGCAACTGTGGCCCATCGACGTGCAGCGCGG	775
D	726	CAAACTGCCCGCAGCGCTGTGGCGACGACCAACTCTCCCTTAGATGGCTTCACTCAA	785
Q	776	TGAGGTGACACGCGCTTACTACCAACACTGCTGTGCGGGCGGGGCGCTGTCTACTCG	835
D	786	TCACATTGGAACAAAGTATTACTTAACTTAAATAACAAAGAAAGATTTCCACTCCG	845
Q	836	ACCAGAGCTCTTCAACGGCGGTCGACGAGACGCGCTGTGAGGACGATACAGCAGCAGG	895
D	846	ACCAGAGCTCTTAAATGAGGTTCTACAGATTCTCAGGTTACGCGACAGCAGCAATC	905
Q	896	CGTGTCTTCAACGCGCACTTCGTGGCAGCCATATTTAGATGGGCAACTTGGGGTGC	955
D	906	AGAACAGCTCTTTATAGACTTTACAGCTGCGCATGTAAAGATGGAAATATTAGCCCTC	965
Q	956	TCACGCGACGCGGAGACGATCAGGCGCACTGGCGG	993
D	966	TCACGTGGCACTAACGGGCAATTCGCCAAAACTGACG	1003

RESULT 3  
US-09-615-192A-368

; Sequence 368, Application US/09615192A  
; Patent No. 6410718

; GENERAL INFORMATION:  
 ; APPLICANT: Bloksberg, Leonard N.  
 ; ADDITIONAL: Haulkaka Tikka

; AFFILIATION: HAWAIIAN LIGNIN  
 ; TITLE OF INVENTION: Materials and Methods for the  
 ; TITLE OF INVENTION: Modification of Plant Lignin Content

FILE REFERENCE: 11000.1003c4U  
CURRENT APPLICATION NUMBER: US/09/615,192A

;; CURRENT FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 08/975,316

; PRIOR FILING DATE: 1997-11-21



```

QY 832 TCGACACGAGACCTTTCACGGCGGGTCCGACGACGCGGTGGAGGACGAGCC 891
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1020 TCTGATCAAGAGCTTTCAATGAGAGTTCGTGTGATTCACGTGTAAGAAATGACAGC 1079
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 892 ACCGCTGCTCTTCACGCGGACCTTGTGTGACGACCATTAAGATGGGCAAGCTGGG 951
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1080 GACACAGGAAAAATTTTCCGAGATTTTCCAGGCAATATGAATAATGAGCAAAATTAAAG 1139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 952 GTGCTACCGGACCGCGGACGACAGATCAGCGCACTGCGGAGTCGTCAAC 1002
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1140 CCCCCAAGAGAGCAATGTCAATAAGAAAAATTGACGAGAAAGTGAAAC 1190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 5

```

US-09-365-150-3
; Sequence 3, Application US/09365150
; Patent No. 6278041
; GENERAL INFORMATION:
; APPLICANT: Lagrimini, Mark
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6278041el Peroxidase Gene Sequences
; FILE REFERENCE: S-31081P1
; CURRENT APPLICATION NUMBER: US/09/365,150
; CURRENT FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic N.
; OTHER INFORMATION: sylvestrin peroxidase gene
US-09-365-150-3

```

```

Query Match      18.88; Score 254.4; DB 4; Length 975;
Best Local Similarity 58.38; Pred. No. 1.6e-44;
Matches 541; Conservative 0; Mismatches 351; Indels 36; Gaps 4;

```

```

QY 111 CCTCTCTGTGTGCGCCACGACAGCTCTGCGGACGTTCTATGAGCTCTCTGCGCCAA 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45 CATCTCGCGCGGACAGCAAGCGCCAGCTGAGCGCCACTTCTAGACACCACTGCCCCAA 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 CCTGCAAGACATCTGTGGGGGGGAGATGACCAAGCGCGTGCAGATGAGAGAGATGG 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 105 CGTGACACGATCGTGGCGGCGGTGATGACAGCGCGACGCGACCGAGCCCGCGCGG 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 231 CGCCTCTGTGCTCAAGGCTCTTCTCAAGACTCTTCTGTTCAAGGCTGCGAGATCGAT 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 165 CCGCAAGATCAATCCCGCTGCACTTCCAGCACTGCTTCTGTAAGCGCTGCGAGCAT 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 291 CCTTCTGACGCGGAGGAGAGAC-----GCGGGCGGAACCTGTAACCTGCTGCG 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 CTTGTGACACGACGACGACGACGACGAGAAAGAGAGCTCTCTGCAAGTGGGCGCGG 284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 345 CGGCTTTAGAGTCATGACACCATCAAGCGGAAGTCGAGGCGCGCGTCCCGCGCTGCT 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 285 CGGCTTTCAGATCGTGGAGACATCAAGACCGCTCTGAGAAAGTGTGCGCGGAGTGT 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 GTCTGCGCGGACATCTCGCGCTTGGCGGCGTGCAGGAAACCACTCTTGGGCGGCGC 464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 345 GAGCTGGCGGACATCTCGCGGCGTGGGAGAGATCGCGGTGTGTGCGCAAGGGCGC 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 465 GACCTGAGAGCTGCGCGCTCGGGGCGGAGCTGAGAGCGGCAAGCGCTGCTGCGCAA 524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 405 TAGCTGGAGAGTGTGTTGGCGCGGCGCAAGAGACCTGACTGCGCAACCGAGCGGCGCAA 464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 525 CAGCAACCGCGCGCGCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 465 CAGCGACATCCCGACGCGCTTTCGAGAGCCTGGCGGTGATGATGCCCACTTACCAACA 524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 585 GGGCTGTGCGCGCGGACATGAGCGGCTGTGTGGGCGGCAACCATGCGGCAAGCGCG 644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 525 GGGCATGAGACCTGACGACGACCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 645 GTGACACGACCTTCCG-----CGCGCGCATCTTCCGCGCGCGCGCGCGCGCGCGCG 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 585 CTGGGCGACCTTCCAGCAGAGAGACTGTTCAACTTCAACGCGACGCGCAACCGCGAC 644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 684 CATCAAGCTCTCTTCCGCGCGCTGCGGACGACGACGACGACGACGACGACGACGACG 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 645 CTTGAGACCGCCACTTCTCTGCAAGCCCTCAAGGACATCTGCGCGCGCGCGCGCGCG 704
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 744 CAACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 CAAGACCTTCAACCAACCTGAGCATGACAGCAACCGCGCGCGCGCGCGCGCGCGCG 764
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 801 CAACCTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 CAACCTTCAAGAGCAACCAAGGCGCTGCTCAAGACGACGAGAGCTGTTCAAGCGCG 824
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 855 CGGCTGCGAGGACGCGCTGTGAGGCAATCAAGCGCGCGCGCGCGCGCGCGCGCGCG 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 CAGCGCGACCATCGCATCTGTAACCGTACCGCGCGCGCGCGCGCGCGCGCGCGCG 884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 915 CTCTGCGACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 CTCTGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 944
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 975 GATCAGCGCGCACTGCGGCTGCTCAAC 1002
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 945 GATCGCGACCGATGCAAGCGCGTGAAC 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 6

```

US-09-365-150-4
; Sequence 4, Application US/09365150
; Patent No. 6278041
; GENERAL INFORMATION:
; APPLICANT: Lagrimini, Mark
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6278041el Peroxidase Gene Sequences
; FILE REFERENCE: S-31081P1
; CURRENT APPLICATION NUMBER: US/09/365,150
; CURRENT FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic N.
; OTHER INFORMATION: tomentisiformis peroxidase gene
US-09-365-150-4

```

```

Query Match      18.18; Score 245; DB 4; Length 975;
Best Local Similarity 58.48; Pred. No. 1.4e-42;
Matches 544; Conservative 0; Mismatches 345; Indels 42; Gaps 5;

```

```

QY 111 CCTCTCTGTGTGCGCCACGACAGCTCTGCGGACGTTCTATGAGCTCTCTGCGCCAA 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45 CATCTCGCGCGGACAGCAAGCGCCAGCTGAGCGCGCACTTCTAGACACCACTGCCCCAA 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 CCTGCAAGACATCTGTGGGGGGGAGATGACCAAGCGCGTGCAGATGAGAGAGATGG 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 105 CGTGACACGATCTGTGGCGGCGTGTGATGACAGCGCGCACGCGCACCGCGCGCGCG 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 231 CGCCTCTGTGCTCAAGGCTCTTCTTCAAGAGCTCTTCTGTTCAAGGCTGCGAGATGAT 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 165 CCGCAAGATCAATCCCGCTTGCATCTTCCAGCACTGCTTCTGTAAGCGCTGCGAGCAT 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 291 CCTTCTGACGCGGAG-----GGAGAAAGACCGCGGCGGAGCACTGAACTGCTGCT 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 CTTGTGACACGACGACGACGACGACGACGAGAAAGAGAGCTGTGCTCCCAAGTGGGCGG 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

OY 342 GCGGCGCTTTGAGTCAATGCAACATCAAGCGAAAGTGCAGCGCCGCTGCCCCGCGCT 401.
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 284 --GGGGCTTCCATCTGTCGACGACATCAAGACCGCCCTCGAGAACGTGTGCCCCGGAGT 341
OY 402 CGTGTGCGCGCCGACATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 461
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 342 GGTGAGCTGGCGCGACATCTGAGCGCTCGGAGACCGAGATCGCGCGTGGCGCGCGAGG 401
OY 462 GCCGACCTGAGAGCTGCGCGCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
    || || || || || || || || || || || || || || || || || || || || || ||
Db 402 CCTAGCTGGAGAGTGTGTTGCGCGCAAGAACCTGACTGCGACCGACGAGAGAG 461
OY 522 CAACAGCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 581
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 462 CAACAGCGCATGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
OY 582 GCAGGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 641
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 522 CAGAGGCGATGACCTGACCGACCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 581
OY 642 CCGGTGACACACTTTC-----CGCGCGCGCATCTACGCGCGAC 680
    || || || || || || || || || || || || || || || || || || || || || ||
Db 582 CCGGTGCGCGACCTTTCGAGCAGAGACTTTCACCTTCAAGCGGAGAGCGGACCGCGACCC 641
OY 681 CGACATCAACGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 740
    || || || || || || || || || || || || || || || || || || || || || ||
Db 642 CACCGTGGACCGCACTTCTCCAGACCTTGCAGGCGCATCTGCCCGCGAGGCGCAACAA 701
OY 741 CCGCAAA--CCTGGCGCGCATTCAGCTGACAGCGCGCGGTGAGGTTTCAGACGCGCTACT 797
    || || || || || || || || || || || || || || || || || || || || || ||
Db 702 CCGCAACACTTTCACCAACCTGACATTCAGCAGCCCGCGCGCGCGCGCGCGCGCGCG 761
OY 798 CACCAACCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 851
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 762 CACCAACCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 821
OY 852 CCGCGGCTGCGAGAGCGCGCTGTGAGGACATACAGCGCGCGCGCGCGCGCGCGCGCGCG 911
    || || || || || || || || || || || || || || || || || || || || || ||
Db 822 CCGCGAGCGCGCACATCGCGCATTCGGAACCGCTACGCGCGCGCGCGCGCGCGCGCG 881
OY 912 CGACTTGTGGCAGCATGATTAGATGGGAGCGTTGGGGTGTCCACCGCGCGCGCGCGCG 971
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 882 CGACTTGTGGCAGCATGATTAGATGGGAGCGCATTCAGTCCCTGACCGCGCGCGCG 941
OY 972 ACAGATCAGGCGCACTGCGCGGTGTCACAC 1002
    || || || || || || || || || || || || || || || || || || || || || ||
Db 942 CGAGATCCGACGACGATGCAAGCGCGGTGAC 972

```

## RESULT 7

```

US-09-615-192A-361
; Sequence 361, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000 100364U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 361
; LENGTH: 916
; TYPE: DNA

```

ORGANISM: *Pinus radiata*  
US-09-615-192A-361

Query Match 17.3%; Score 234; DB 4; Length: 916;  
Best Local Similarity 58.7%; Pred. No. 2.7e-40;  
Matches 452; Conservative 0; Mismatches 300; Indels 18; Gaps 2;

```

OY 143 CCAGTCTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 134 CCAGTCTATGATTAATCTTGTTCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 193
OY 203 AGCGCTGCGCAAGTGAACAGAGATGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTG 262
    || || || || || || || || || || || || || || || || || || || || || ||
Db 194 AAGCAGTGTAAAGAAAGAAAGTATGGGGGCGATCGCTTCCGCTTATTTCCACACT 253
OY 263 GCTTGTGTAAGGCTGAGGAGATCGATTCCTGTA-----CGCGAGAGGG 310
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 254 GCTTGTGTAAGGCTGAGGAGATCGATTCCTGTA-----CGCGAGAGGG 313
OY 311 AGAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 370
    || || || || || || || || || || || || || || || || || || || || || ||
Db 314 AGAATATGCTCTTCCCAATTAACATTCGCGAGGGGTTGAGGTGATGATGATGATG 373
OY 371 AGCGGAACGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 430
    || || || || || || || || || || || || || || || || || || || || || ||
Db 374 AGAGCAACTGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 433
OY 431 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 490
    || || || || || || || || || || || || || || || || || || || || || ||
Db 434 CTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 493
OY 491 GGGACTGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 550
    || || || || || || || || || || || || || || || || || || || || || ||
Db 494 GAGATTCAGCAAGACGAGCATTAAGGATGAGCAAGATTAATTCGCGCTCCACTTCCA 553
OY 551 GCGTGGCGACGCTATCTCCCTGTTGCGGAGAGGCGCTGCGCGCGCGCGCGCGCGCG 610
    || || || || || || || || || || || || || || || || || || || || || ||
Db 554 ATCTGACGAACTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 613
OY 611 CGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 670
    || || || || || || || || || || || || || || || || || || || || || ||
Db 614 CACTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 673
OY 671 AGCGGACACCGCATATCAAGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 730
    || || || || || || || || || || || || || || || || || || || || || ||
Db 674 ACAGGACACCGCATATCAAGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 733
OY 731 C-----CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 784
    || || || || || || || || || || || || || || || || || || || || || ||
Db 734 CCAGAGGCTCCGAGACGAGCACTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 793
OY 785 ACAGGCGCTACTTACCAACCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 844
    || || || || || || || || || || || || || || || || || || || || || ||
Db 794 ACAGGCGCTACTTACCAACCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 853
OY 845 TCTTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 894
    ||||||| || || || || || || || || || || || || || || || || || || ||
Db 854 TCTTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 903

```

## RESULT 8

```

US-09-097-319A-5
; Sequence 5, Application US/09097319A
; Patent No. 6384207
; GENERAL INFORMATION:
; APPLICANT: Ainley, Michael
; APPLICANT: Armstrong, Katherine
; APPLICANT: Belmar, Scott
; APPLICANT: Folkerts, Otto
; APPLICANT: Hopkins, Nicole
; APPLICANT: Menke, Michael A.
; APPLICANT: Paredy, Dayakar
; APPLICANT: Petolino, Joseph F.

```







Db 181 GACATATGCGAAGGCGAACTGCCCCAGCACTGGTCGACCAACACCTCTCCCTTG 240  
QY 760 GACGTGCACACCGCGGTGAGTTCACACAGGCGCTACTTACCAACCTGCTCTCGGCGG 819  
Db 241 GAGTGTCTACCCCTACCACTTATTAACAACAAGTATTAACCTGTAAGAAAGCCAAAG 300  
QY 820 GGCCTGTTCCACTCGACACGAGCTTTCAACGCGGGTGCAGAGCGCGCTGTGAGG 879  
Db 301 GGACCTTTCACCTCGATCAGAGCTATTTATGAGGTTCCACAGACTCTAGAGTTAGT 360  
QY 880 CAGTACAGGCGCAAGCGCTCTGCTTTCACAGCCGCTGCGGACCTATTAAGATG 939  
Db 361 ATCTACAGACCAAGTCAACCCATTTTCTTACTGACTTTCAGAGCGCGCTGTAATG 420  
QY 940 GGCACGTTGGGCTGCTACCGGACCGCGACAGATCAGACCACTGCGGGTGTG 999  
Db 421 GGTAAATATTAGTCCCTCCTCAGTGCACCAAGCGGAGATCCGCAAACTGCAGGAAGTC 480  
QY 1000 AA 1001  
Db 481 AA 482

RESULT 10  
US-09-615-192A-365  
Sequence 365, Application US/09615192A  
Patent No. 6410718  
GENERAL INFORMATION:  
APPLICANT: Bloksberg, Leonard N.  
APPLICANT: Havukkala, Ilkka  
TITLE OF INVENTION: Materials and Methods for the  
FILE REFERENCE: 11000.1003c4U  
CURRENT APPLICATION NUMBER: US/09/615,192A  
CURRENT FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 08/975,316  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11  
PRIOR APPLICATION NUMBER: US 09/169,789  
NUMBER OF SEQ ID NOS: 405  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 365  
LENGTH: 646  
TYPE: DNA  
ORGANISM: Pinus radiata  
US-09-615-192A-365

Query Match 13.0%; Score 175.8; DB 4; Length 646;  
Best Local Similarity 58.6%; Pred. No. 3.4e-28;  
Matches 331; Conservative 0; Mismatches 222; Indels 12; Gaps 1;

QY 98 TCCCGCTTCCCTCCTCTGTCGTCGACAGCTCCGCCAGCTTATGCT 157  
Db 81 TCCCTCTGCTCTTTTTCGACCGCTTGCTTCTCACTCACTCACTATTAATA 140  
QY 158 CCTCTGCCCAACCTGACAGCATCTTCGCGGCGCATGACCCAGCCCTGCAAGT 217  
Db 141 CGTCGTGTCCCAAACTCTGGAACGCTGAGGCTGACATGAAAGCGGCTGCAATG 200  
QY 218 AGAAGAGAGGAGGCGCTCTGTCGTCAGGCTTCTTCCAGACAGCTTCGTAAGCT 277  
Db 201 AGAAGAGCATGAGGCGCATATGCTCTGCTTCACTTATGATGTTTCTGCAATG 260  
QY 278 GCGACGATCGATCTTCTGAGCGC-----GAGGAGGAGAAGACCGCGGCG 325  
Db 261 GCGATGGGTAGTGTGTTGAGACACTTTCAGTCTAAGTGGGAAAAAGACTCTCT 320  
QY 326 CGAAGCTGAAGTGGTGGCGCTTTGAGTATGACACCACTCAAGGGGAACGTCAGG 385  
Db 321 CCAAGCAACAATTCGTTAGGGGTTTCGACGCTATGACACCACTCAAACTCAAGTGAAG 380

QY 386 CCGCGTCCCGCGCGCTGTCGTCGCGGACATCTGCGCTTGGCGCGGACGGA 445  
Db 381 CAGTTTGCAGCGGATTCGATGCTGCTGCTGATCTTTGCTATTAAGCTAGAGATTCTG 440  
QY 446 CCAACTTTCGCGGCGCGGACCTGAGCGTCCGCTGCGGCGGAGCTGACGAGG 505  
Db 441 TCGTCGAATTGGGAGGACCAACATGAGACAGTCTGCTTGAAGAGAGACCTCAACTG 500  
QY 506 CCAAGCGCTCTGTCGCAACAGCAACCCCGCGCGAGCGGACGCTGAGGCTCA 565  
Db 501 CCAAGCTTAAGCGCGGCAACCAACCACTTCCGCTCCACTTCAATCTCACTGCTCA 560  
QY 566 TCTCCCTGTTGCGGACGAGGCGCTGTCGCGCGGACATGACGCGCTGTGGCGCG 625  
Db 561 TCTCATCTTTCAAGCAGAGCGGCTTTCACCAAGATATGATGTTCTTATCAGGTGAC 620  
QY 626 ACACCATGCGGCGGCGGCTGAC 650  
Db 621 ATTCATGCGGCGGCGGCTGATGAC 645

RESULT 11  
US-09-615-192A-359  
Sequence 359, Application US/09615192A  
Patent No. 6410718  
GENERAL INFORMATION:  
APPLICANT: Bloksberg, Leonard N.  
APPLICANT: Havukkala, Ilkka  
TITLE OF INVENTION: Materials and Methods for the  
FILE REFERENCE: 11000.1003c4U  
CURRENT APPLICATION NUMBER: US/09/615,192A  
CURRENT FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 08/975,316  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11  
PRIOR APPLICATION NUMBER: US 09/169,789  
NUMBER OF SEQ ID NOS: 405  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 359  
LENGTH: 659  
TYPE: DNA  
ORGANISM: Pinus radiata  
US-09-615-192A-359

Query Match 9.0%; Score 121.6; DB 4; Length 659;  
Best Local Similarity 54.6%; Pred. No. 6.8e-17;  
Matches 272; Conservative 0; Mismatches 214; Indels 12; Gaps 1;

QY 151 TATGCTCTCTCTGCCCCAAGCTGACAGCATCTTCGCGGCGGATGACCCAGCGTC 210  
Db 156 TATCAGAAAGCTCTGCAAAATGCCAGGCTATCATTAATGCTGATTCAGATGCTG 215  
QY 211 GCAAGTACAGAGATGAGGCGCTCTGCTCAGGCTCTTCTTCCAGATGCTGCT 270  
Db 216 AAAAAGGAGCGAATGCTCTCTCTGCTCTGCTGCAATTTTCAGATGCTGTTGTT 275  
QY 271 CAAGGCTGCGAGGATGATCTCTTCTGACGCGG-----AGGGAAGAAC 318  
Db 276 CAGGCTGTGATGCTTCAATCTGCTGATGACACTGCTAGTTTCACAGGGGAGAACCA 335  
QY 319 GCGGCGCGAAGCTGATCTGCTGCGGCTTTGAGTCACTGACACCATCAAGCGAAC 378  
Db 336 GCAATTTCTTAACGAATCTGTAAAGGCTTTGAGTGAAGATTAAGATTAAGCAAA 395  
QY 379 GTCGAGCGCGGCGCGCGCTGCTGCTGCGGACATCTGCGGCTTGGCGCGGCG 438  
Db 396 TTGAGAGAAAGCATGCTCTGAGTGTGCTCATGTGCTGATCTTCTGCTGCGGACCGCT 455  
QY 439 GACGGAACCAACTCTCTGCGGCGGCGGACCTGGAAGCTGCGGCTGCGGCGGAGCTCG 498



ADDRESSEE: BANNER & ALLEGRETTI, LTD.  
 STREET: 10 S. WACKER DRIVE, SUITE 3000  
 CITY: CHICAGO  
 STATE: ILLINOIS  
 COUNTRY: U.S.A.  
 ZIP: 60606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/462,695  
 FILING DATE: 5-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/190,029  
 FILING DATE: 28-FEB-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/GB92/01389  
 FILING DATE: 27-JUL-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GB 9202401.7  
 FILING DATE: 05-FEB-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GB 9116325.3  
 FILING DATE: 28-JUL-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: JOHN J. McDONNELL  
 REGISTRATION NUMBER: 26,949  
 REFERENCE/DOCKET NUMBER: 94,062-A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312-715-1000  
 TELEFAX: 312-715-1234  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1072 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 11..1057  
 OTHER INFORMATION: /codon\_start=11  
 OTHER INFORMATION: /function="Gene for IGF-II/HRP fusion protein"  
 OTHER INFORMATION: /product="Synthetic DNA"  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 1..6  
 OTHER INFORMATION: /function="HindIII cleavage site"  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 1067..1072  
 OTHER INFORMATION: /function="EcoRI cleavage site"  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 11..856  
 OTHER INFORMATION: /function="HRP coding sequence"  
 OTHER INFORMATION: (fragment)  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 857..1057  
 OTHER INFORMATION: /function="IGF-II coding sequence"  
 US-08-462-695-9

Query Match 8.8%; Score 119.6; DB 2; Length 1072;  
 Best Local Similarity 53.5%; Pred. No. 2e-16;  
 Matches 304; Conservative 0; Mismatches 249; Indels 15; Gaps 2;  
 133 CACCTCTCCGCCAGCTTATGCTCTCTGCCCCAACCTGACAGACATGTTGGGCG 192

Db 14 CAGTTAACCCCTCATCTTACGACAAATAGCTGTCCCAAGTGTCCAAATGTTCCGAC 73  
 193 GCGATGACCCGACCCGCTCGCAAGTACGAGAGATGGGCGCTCTCTGCTAGGCTTTC 252  
 Db 74 ACAATGCTCAACGAGCTCAGATCCGATCCAGATCCAGATCCGATCCAGATCCAG 133  
 253 TTCACAGCTGCTCTGTTCAAGGCTGCGAGGATGATGATCTCTGAC----- 300  
 Db 134 TTCATGATGCTGCTGGAATGGTTCGACGCTAGCATTTATGGAACAACACCACT 193  
 301 GCCGAGGAGGAGAGAACCCGCGGACCTGAACTGCTGCGGCGCTTGAAGTATC 360  
 Db 194 TTCGACACTGAAAGATGATCGGGAACGCTAACAGGCGCAGGCGCTTCCAGTATC 253  
 361 GACACATGCAAGGAGAAAGCTGCGGAGGCGGCTCCCGGCTGCTGCTGCGGAGATC 420  
 Db 254 GATGCAATGAGGCTGCGCTGTTGATGATGATCCCAAGACAGATGATGATGATG 313  
 421 CTCGCGCTTCCGCGCGCGGACCAACCTTCTCGGCGGCGGCGGCTGCGGCGG 480  
 Db 314 CTGACTATGCTGCGGACAGAGCTGACTCTTGCAGGCGGACGCTCTGAGAGTCCG 373  
 481 CTCGGGGGGGGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 Db 374 CTCGGTGAAGTGAATCTCTACAGAGGATTCATGATCTGAGAGAGAGAGAGAG 433  
 541 CCGAGGCGCAGCTCGGACAGCTCATCTCTCTGCGAGGAGGAGGCTGCTGCG---CG 597  
 Db 434 CCATTCTCACCTCGCCCGCCAGCTGAAGATGATCTTGAAGAGTGTGATTCGCTCG 493  
 598 CGGCAATGAGGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 657  
 Db 494 AGTACCTTGTGGCTCTGTCCGAGGACACACATTTGGAAAGACAGTATGAGTTATC 553  
 658 CGGCGCCGATCTACGCGGACACCGACA 685  
 Db 554 ATGATAGGCTCTACAAATTTACAGACA 581

RESULT 14  
 US-09-615-192A-371  
 Sequence 371, Application US/09615192A  
 Patent No. 6410718  
 GENERAL INFORMATION:  
 APPLICANT: Bloksberg, Leonard N.  
 APPLICANT: Havukkala, Ilkka  
 TITLE OF INVENTION: Materials and Methods for the  
 TITLE OF INVENTION: Modification of Plant Lignin Content  
 FILE REFERENCE: 11000.1003c4U  
 CURRENT APPLICATION NUMBER: US/09/615,192A  
 CURRENT FILING DATE: 2000-07-12  
 PRIOR APPLICATION NUMBER: US 08/975,316  
 PRIOR FILING DATE: 1997-11-21  
 PRIOR APPLICATION NUMBER: US 08/713,000  
 PRIOR FILING DATE: 1996-09-11  
 PRIOR APPLICATION NUMBER: US 09/169,789  
 PRIOR FILING DATE: 1998-10-09  
 NUMBER OF SEQ ID NOS: 405  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 371  
 LENGTH: 1522  
 TYPE: DNA  
 ORGANISM: Pinus radiata  
 US-09-615-192A-371

Query Match 8.8%; Score 119.2; DB 4; Length 1522;  
 Best Local Similarity 49.4%; Pred. No. 2.6e-16;  
 Matches 443; Conservative 0; Mismatches 423; Indels 30; Gaps 4;  
 136 CTCCTGCGCCAGCTTATGCTCTCTGCCCCAACCTGACAGACATGTTGGGCGG 195  
 Db 182 CTTTCTTGACGCTTACAGCTCAGATTCGCGCTCTTGGAGTGCATAGTGTGAGACGC 241

```

QY 196 ATGACCGAGCGCGCGGACAGTGGAGGATGGGCGGCTCTGCTGCAAGCTCTTCTC 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 ATGGAAGCTTATTTAGTGCAGACATCACAGAGCTGCAGATTGTGAGCTCAGCTTC 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 CACGACTGCTTCTTCAAGCTGCGAGGATCGATTCCTTCTGCAGCCGG--AGGGAG 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 CAGACTGCTTGTCTCAGAGATGAGATGGGTGCGTGTGTTGAACGACAGTGAAGTGA 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 313 AAGACCGCGGCGGCAACTGAA--CTCGGTGCGGCGCTTGAAGTCATGCACACATTC 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 CAAGGCGTCCCCCAACTTATCACTCAGACGCGAGCTTAAAGATTATTAAGCATTC 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 AAGCGAAGCTGAGAGCGCGCTGCGCGCTGCTGCTGCTGCGCGACATCTCGGCTT 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 AAGAGAACGTGAGAGCGCGCTGAGGAAATGTGTGCTGCTGCGACATGTACTTAA 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 430 GCGGCGCGGAGGAGAACCACTTCTGCGCGCGCGCGACGTGAGCGCGCTCGGGCGG 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 GCGAGCTGTGACTCGCTGTATGCTGAGAGACCGCTTCTACCCCTTACACTGCGCGC 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 490 CGGGA---CTGACGACGCGCGACGCGCTGCTGCTGCGCAACGACACCCCGCGGAG 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 AAGGACAGCTTACCTTCCCAATGACATGACACGCTTCTGCGCAATTGCGATCCCAAC 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 547 GCGAGCTGCGGACGCTGCTCCTGCTGCTGCGGAGGAGCGCGCTGCGCGCGGAGT 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 TCCATGTATACGCGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 607 ACAGCGCTGCTGCGGCGCGCACACATCGGCGAGCGCGGTCACACCTTCCGCGCGC 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 GTGCGCTTCTAGAGAGACATCAATTTGGAGAGCAACTGCTCTCTGCGACACAGA 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 667 ATCTAGGAGAACGACGATCAAGCCCTCTGCGGCGCTGCGGCGGAGAGAGTGGCC- 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 CTATATACAGACACACCGGTACAAATGCGGAGATCCACGATGAGACGAGATTCGCT 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 726 -----GCGGTGCGGCGGCGGCAACCTGCGGCCCATCGACGTG 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 782 AAGAACTTATCTACCTGCTTACAGTACACGCTTAAACACCAAAATGGATATTT 841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 CAGAGCGCGGTAGGTGACACGCGCTTCTACCAACCTGCTGCGGCGGCGCTG 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 CGACCTCCAAATGTGTGACACAAATACATCTGCTGCTGCTGCTGCTGCTGCTGCTG 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 TTCACCTGCGACGAGAGCTTCTCAAGCGCGGTGCGAGAGCGGCTGAGGACGTAC 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 TTCACCTGCGACGAGAGCTTCTCAAGCGCGGTGCGAGAGCGGCTGAGGACGTAC 961
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 886 AGCGGCGAGCGCTGCTTCAAGCGCGCTTCTGCGGAGCGATGATGAGGAGGAC 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 962 GCGGAGATCAAGAGCTCTCTTTCAGACATTTGTGCTGAGCATGTCCAAATGGGCG 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 946 GTTGGGCTCTACCGGACCGCGGAGATCAAGCGGACATGCGGCGCTGCTCA 1001
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1022 CTGGATGTCTCACAGAGGAGGAGATCCGTAAGAACTGCGGCTGCGAA 1077
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 15

```

US-09-615-192A-370
; Sequence 370, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukala, Ilka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c4u
; CURRENT APPLICATION NUMBER: US/09/615.192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975.316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713.000

```

```

; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 370
; LENGTH: 1073
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-370

```

```

Query Match 8.3%; Score 113; DB 4; Length 1073;
Best Local Similarity 51.0%; Freq. No. 4.7e-15;
Matches 340; Conservative 0; Mismatches 300; Indels 27; Gaps 2;

```

```

QY 367 ATCAACGAGAGTGCAGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 ATTAAGAGTGCACCTTGAAGAGAGTGCACCAAACTGATGCTGAGATTTCTCGCT 65
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 427 CTTCGCGCGCGGAGGAGAACCACTTCTGCGCGCGCGGAGACTGAGAGCTGCGCTG 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 ATTGCATCTGCTGATTCAGTGTCTGAGTGAAGGCTGCGGAGGATTTACTGAGG 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 487 CCGCGGAGTGCAGACGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 AGGAGAGATTGCAAGAGTGCAGATTGAGTGAAGTGCACCAACATATCCGCGCGCAC 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 547 GCGAGCTGCGGACGCTCATCTCTGCTGCGGAGGAGCGCGCTGCGCGCGGAGT 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 TCAACTCTGACAGCGCTTCTACACCAAGTTCAAACTACAGGCTGAGTAGAGTAC 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 607 ACAGCGCTGCTGCGGCGCGCACACATCGGAGAGCGCGGTCACACCTTCCGCGCGC 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GTATCCCTTTCAGGAGATGACACATCGGCTATCTGATGATCAAAATTTACAGGAG 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 667 ATCTAGGCG-----GACACCGACATCAACGCGCTCTCTGCGGCG 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 CTTCACACACAGAGTGAAGTGAAGTGGCTGCCAGATCTTCACTTAACAGGGTTACTATGCT 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 CTGCGGACAGACGAGTGCAGCGCGGTGCGGCGGAGGAGCACTGCGGCCATGACGTG 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 CGGCTGAATTCGAGATGTCACAAATCTGAGAGAGATATTAATCTGTCATTTGGAATTC 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 CAGAGCGGCTGAGTGTGACACGCGCTTCTACCAACCTGCTGCTGCGGCGGCGCTG 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 GTGACTCTACCAAAATTCATCTACTACTACTACTACTACTACTACTACTACTACTACT 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 TTCACCTGCGACGAGAGCTTCTCAAGCGCGGTGCGAGAGCGGCTGAGGAGG 879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 TTGAACACAGACAGAAATTTGTCGCAAGAGGCTGAGGAGAGGAGGAGGAGTAA 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 880 CAGTACAGGCGCGAGCGCTGCTTCAAGCGCGCTTCTGCGGAGCGCATGATGAGATG 939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 CTTTATGACAGCAATGAGAGAGCTTCTTCTCAAAAGTGTGATATATATATGAGAGATG 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 940 GCGAAGTGTGGGTGCTCAAGCGCGCGGAGACATGAGGCGCAATCCGCGGTGCTC 999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 GGAACATCAAGCTCTTACAGGACAGTGGAGAGAAATCAGGCTCACTGCTAAGATT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1000 AACAGCT 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 666 AACAGTT 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Search completed: May 3, 2003, 09:25:44
Job time : 57 secs

```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 10:03:52 ; Search time 1455 Seconds

(without alignments)  
15071.264 Million cell updates/sec

Title: US-10-047-825-3

Perfect score: 1354

Sequence: 1 aattcgccagcagcgttaacg.....aaaaaaaaaaaaaaaaaaaaa 1354

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 16

Total number of hits satisfying chosen parameters: 1328565

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:\*  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_estc2:\*  
11: gb\_hic:\*  
12: gb\_estc3:\*  
13: gb\_estc4:\*  
14: gb\_estc5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_iny:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	441	32.6	848	9	AA979912 MEST3-D6
2	270	19.9	446	12	BG841858 MEST7-D0
3	241	17.8	538	12	BG842311 MEST9-D0
4	166	12.3	220	14	T18410 6C02D10C7 e
5	92	6.8	850	9	A1374530 MEST3-D6
6	85	6.3	383	13	BW501068 PAC000000

Result	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	AA979912	MEST3-D6	848 bp	linear	EST	Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC	Wen, T.J., Ashlock, D.A. and Schnable, P.S.	Expressed Sequence Tags from B73 Maize Seedlings	Unpublished (1997)	Contact: Schnable, PS	Schnable Laboratory Iowa State University G405 Agronomy, Ames, IA 50011, USA Tel: (515)-294-0975 Fax: (515)-294-2299 Email: schnable@iastate.edu
2	AA979912	MEST3-D6	446 bp	linear	EST	Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC	Wen, T.J., Ashlock, D.A. and Schnable, P.S.	Expressed Sequence Tags from B73 Maize Seedlings	Unpublished (1997)	Contact: Schnable, PS	Schnable Laboratory Iowa State University G405 Agronomy, Ames, IA 50011, USA Tel: (515)-294-0975 Fax: (515)-294-2299 Email: schnable@iastate.edu
3	AA979912	MEST3-D6	538 bp	linear	EST	Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC	Wen, T.J., Ashlock, D.A. and Schnable, P.S.	Expressed Sequence Tags from B73 Maize Seedlings	Unpublished (1997)	Contact: Schnable, PS	Schnable Laboratory Iowa State University G405 Agronomy, Ames, IA 50011, USA Tel: (515)-294-0975 Fax: (515)-294-2299 Email: schnable@iastate.edu
4	T18410	6C02D10C7 e	220 bp	linear	EST	Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC	Wen, T.J., Ashlock, D.A. and Schnable, P.S.	Expressed Sequence Tags from B73 Maize Seedlings	Unpublished (1997)	Contact: Schnable, PS	Schnable Laboratory Iowa State University G405 Agronomy, Ames, IA 50011, USA Tel: (515)-294-0975 Fax: (515)-294-2299 Email: schnable@iastate.edu
5	A1374530	MEST3-D6	850 bp	linear	EST	Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC	Wen, T.J., Ashlock, D.A. and Schnable, P.S.	Expressed Sequence Tags from B73 Maize Seedlings	Unpublished (1997)	Contact: Schnable, PS	Schnable Laboratory Iowa State University G405 Agronomy, Ames, IA 50011, USA Tel: (515)-294-0975 Fax: (515)-294-2299 Email: schnable@iastate.edu
6	BW501068	PAC000000	383 bp	linear	EST	Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC	Wen, T.J., Ashlock, D.A. and Schnable, P.S.	Expressed Sequence Tags from B73 Maize Seedlings	Unpublished (1997)	Contact: Schnable, PS	Schnable Laboratory Iowa State University G405 Agronomy, Ames, IA 50011, USA Tel: (515)-294-0975 Fax: (515)-294-2299 Email: schnable@iastate.edu

## ALIGNMENTS

AA979912 MEST3-D6, TW1412, Seq ISDN2 Zea mays CDNA clone MEST3-D6 5', mRNA  
sequence.  
AA979912  
AA979912.1 GI:3157290  
EST  
Zea mays.  
Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC  
Clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 848)  
Wen, T.J., Ashlock, D.A. and Schnable, P.S.  
Expressed Sequence Tags from B73 Maize Seedlings  
Unpublished (1997)  
Contact: Schnable, PS  
Schnable Laboratory  
Iowa State University  
G405 Agronomy, Ames, IA 50011, USA  
Tel: (515)-294-0975  
Fax: (515)-294-2299  
Email: schnable@iastate.edu  
PCR primers  
FORWARD: tw1412 (5'-GAAGATACCCACCAACC-3')  
BACKWARD: tw1412 (3'-TAATACGACCTCATTAGGAC-3')  
Plate: MEST3 Row: D Column: 6  
Seq primer: tw1412 (5'-GAAGATACCCACCAACC-3').



Db	123	AAGAGCAACCTCCGCGCCCGGACGGCCAGCTCCGACAGCTCACTCTCCCTGTTCCGCAGG	64
QY	583	CAGGCGCTGTCGCGCGGCGGACATGACGGCGGTGTGCGGGCGGACACATCATGGGCAAGGC	642
Db	63	CAGGGCTCTTCGCGCGGACATGACGGCGGTGTGCGGGCGGACACATCATGGGCAAGCC	4
QY	643	CGG	645
Db	3	CGG	1

RESULT 3	538 bp	MRNA	linear	EST 29-MAY-2001
LOCUS				
BG842311				
DEFINITION	ME829-D08.T3 ISUM4-TN Zea mays CDNA clone			
ACCESSION	BG842311			
VERSION	sequence.			
KEYWORDS	BG842311.2			
SOURCE	GI:14244343			
ORGANISM	Zea mays.			
	Zea mays.			

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 538)  
Qiu,F., Cui,F., Guo,L., Ashlock,D.A, Wen,T.J. and Schnable,P.S  
Expressed Sequence Tags from B73 Maize Seedlings and SILs  
Unpublished (2001)  
On May 25, 2001 this sequence version replaced gi:14208633.

Iowa State University  
4405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA  
Tel.: 515-294-0975  
Fax: 515-294-2299  
Email: schnable@iastate.edu

PCR Primers

FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)  
BACKWARD: T3 (ATT AAC CCT CAC TAA AG)  
Seq primer: PRIMER T3 (ATT AAC CCT CAC TAA AG).

```

FEATURES
source
1. 538
location/Qualifiers
1. 538
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/cloae="MEST25-D08"
/clone_11b="ISUM4-ITN"
/tissue_type="Seedling and silk"
/lab_host="DH10B"
/note="Vector: pT73PAC; Site.1: EcoRI; Site.2: NotI;
ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'
AAGTGAAGATTCGGCGCCAGAGATATTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT73PAC vector. The library
then went through one round of normalization to COT value
of 5 based on the methods of Marcello Bento Soares (Genome
Research 6: 791-806, 1996)."

```

	Query Match	17.8%	Score 241	DB 12	Length 538
	Best Local Similarity	99.7%	Pred. No. 0		
	Matches 291	Conservative	0	Mismatches 1	Indels 0
				Gaps 0	
Qy	677	ACACCGCATCAACGGCTCTTCGCGGCGCTGCGGCAACAGACGTGCGCGCGGTCCGGCG	736		
Db	1	ACACCGCATCAACGGCTCTTCGCGGCGCTGCGGCAACAGACGTGCGCGCGGTCCGGCG	60		

QY	797	TCACCAACCTGCTGTCGCGCGCGGGGCGCTGTTCACACTCGAACGAGACGCTCTTTCACACGGCG	856
Db	61	GGCAGCGGCAACCTGGCGGCCCATTCACAGCTGCAGACGGCGGGTGGAGTTTCACACAGCGCTACT	120
QY	737	GGCAGCGGCAACCTGGCGGCCCATTCACAGCTGCAGACGGCGGGTGGAGTTTCACACAGCGCTACT	796
Db	61	GGCAGCGGCAACCTGGCGGCCCATTCACAGCTGCAGACGGCGGGTGGAGTTTCACACAGCGCTACT	120
QY	797	TCACCAACCTGCTGTCGCGCGCGGGGCGCTGTTCACACTCGAACGAGACGCTCTTTCACACGGCG	856
Db	121	TCACCAACCTGCTGTCGCGCGAGGGGCGCTGTTCACACTCGAACGAGACGCTCTTTCACACGGCG	180
QY	857	GGTCGAGAGACCGCGTGGCTGAGGCGCATGCACGCCCGACGCGCTCGCGCTTTCACACGGCGACT	916
Db	181	GGTCGAGAGACCGCGTGGCTGAGGCGCATGCACGCCCGACGCGCTCGCGCTTTCACACGGCGACT	240
QY	917	TCCTGGCAGCCATGATTTAGAGATTGGGCAACAGTTGGGGTGCACACGGGCAACCGC	968
Db	241	TCCTGGCAGCCATGATTTAGAGATTGGGCAACAGTTGGGGTGCACACGGGCAACCGC	292

RESULT 4	230 bp	MRNA	linear	EST 17-OCT-1996
LOCUS	118410	6C0210L7	Ze mays cDNA clone 6C0210.5' end	
DEFINITION	118410	similar to etiolated seedling Zea mays peroxidase, mRNA sequence.		
ACCESSION	118410			
VERSION	118410.1	GI:474233		
KEYWORDS	EST			
SOURCE	Ze mays.			
ORGANISM	Ze mays.			

REFERENCE	1 (bases 1 to 230)
AUTHORS	Shen,B., Carretero,N., Torres-Jerez,I., Stevenson,R., McCreery,T., Hendricks,T., Baysdorfer,C., Almira,E., Ferl,R., Habbem,J. and Larkin,B.
TITLE	Partial sequencing and mapping of clones from two maize cDNA libraries
JOURNAL	Plant Mol. Biol. 26, 1085-1101 (1994)
MEDLINE	9511093
COMMENT	Contact: The maize cDNA Project

Helentjaris TG (primary contact)  
Dept. of Plant Sciences  
University of Arizona  
Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721  
ph: 602-6218-746  
fax: 602-621-7186  
E-mail: helnjars@ccit.arizona.edu

Chris Baysdorfer  
Department of Biological Sciences, School of Science  
California State University, Hayward  
Hayward, CA 94542  
ph: 510-881-3459  
fax: 510-727-2035  
E-mail: cbaysdore@sl.csu Hayward.edu

Rob Ferl  
Interdisciplinary Center for Biotechnology Research  
DNA Sequencing Core  
University of Florida  
P.O. Box 100695  
Gainesville, FL 32611-0695  
ph: 904-392-1928, ext. 301  
fax: 904-392-4072  
E-mail: roberf@nemr.ncrdc.ufl.edu  
Seq primer: T7  
Location/Qualifiers  
1. 230

```

- FEATURES
  SOURCE
    Location/Qualifiers
      1..230
        /organism="Zea mays"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="6C02d10"

```



/clone\_11b="etiolated seedling"  
 /lab\_host="DH10B"  
 /note="vector: ziploxx; site\_1: SalI; site\_2: NotI;  
 oligo-dt selected mRNA was prepared from above-ground  
 tissue from 10 day-old etiolated seedlings. ds-cDNA was  
 prepared by priming with a NotI oligo-dt oligomer and the  
 second strand by RNase-nicking of the DNA:RNA hybrid with  
 DNA PolI fill-in. Sali adaptors were added to the ends,  
 the ds-cDNAs were then digested with NotI and  
 size-selected. These were directionally-cloned into the  
 ziploxx phage vector, excised as plasmids, and then  
 individually analyzed."

BASE COUNT 37 a 82 c 60 g 48 t 3 others  
 ORIGIN

Query Match 12.3%; Score 166; DB 14; Length 230;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 GCTTAAGCAAGTAGCTTCATTACCCAGAGCGAGGACAGAGGAGCGCTTGCCATGGCG 72  
 DB 4 GCTTAAGCAAGTAGCTTCATTACCCAGAGCGAGGACAGAGGAGCGCTTGCCATGGCG 63

OY 73 TCTCCACCTGATGCAATGCTGCTGCGCTTCCTCCCTCCTGCTGCGCCACGCA 132  
 DB 64 TCTCCACCTGATGCAATGCTGCTGCGCTTCCTCCCTCCTGCTGCGCCACGCA 123

OY 133 CACCTCTGCGCCAGCTTCATAGCTCTCTCTGCTGCGCCACCTGCA 178  
 DB 124 CACCTCTGCGCCAGCTTCATAGCTCTCTCTGCTGCGCCACCTGCA 169

RESULT 5  
 AI374530/C 850 bp mRNA linear EST 21-JAN-1999  
 LOCUS MEST3-D6.POLYT-N.Seq ISUM2 zea mays cDNA clone MEST3-D6.5', mRNA  
 DEFINITION sequence.  
 ACCESSION AI374530  
 VERSION AI374530.1 GI:4174550  
 KEYWORDS EST.  
 SOURCE zea mays.  
 ORGANISM zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 850)  
 Men,T.J., Ashlock,D.A. and Schnable,P.S.  
 Expressed Sequence Tags from B73 Maize Seedlings  
 Unpublished (1997)  
 Contact: Schnable, PS.  
 Schnable Laboratory  
 Iowa State University  
 6405 Agronomy, Ames, IA 50011, USA  
 Tel: (515)-294-0975  
 Fax: (515)-294-2299  
 Email: schnable@iastate.edu  
 PCR Primers  
 FORWARD: CW1412 (5'-GAAGATACCCCAACCAACC-3')  
 BACKWARD: T7-YJ (5'-TAATACGACTCACTAAGGCG-3')  
 Plate: MEST3 row: D column: 6  
 Seq primer: POLYT-N (5'-TTTTTTTTTTTTTTTTTTTTTTT(AGC)-3').  
 Location/Qualifiers  
 1..850  
 /organism="Zea mays"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="MEST3-D6"  
 /clone\_11b="ISUM2"  
 /tissue\_type="above ground tissues"  
 /dev\_stage="Two-leaf-stage green seedling"  
 /lab\_host="XLI-MPR Blue"  
 /note="Organ: green seedlings; Vector: PAD-GAL4; Site 1:  
 EcoRI; Site 2: XhoI; ds-cDNA molecules were generated as

follows. First-strand cDNA was prepared from oligo-dT  
 selected mRNA by priming with an XhoI oligo-dT primer. The  
 resulting DNA:RNA hybrid was treated with RNase H and used  
 as a template for DNA PolI-catalyzed second strand  
 synthesis. After the addition of EcoRI adaptors, the  
 ds-cDNAs were digested with XhoI and size-selected. The  
 resulting molecules were directionally cloned into the  
 EcoRI and XhoI sites of the hybridZAP lambda vector  
 (Stratagene) and excised as PAD-GAL4 plasmids."

BASE COUNT 168 a 226 c 210 g 194 t 52 others  
 ORIGIN

Query Match 6.8%; Score 92; DB 9; Length 850;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 877 AGGCAATACAGGCGGCGGCTGCTTCAAGCCGACATCTGACACCAATGATTAG 936  
 DB 398 AGGCAATACAGGCGGCGGCTGCTTCAAGCCGACATCTGACACCAATGATTAG 339

OY 937 ATGGGCAAGCTGGGGTCTCACCAGCGACCGC 968  
 DB 338 ATGGGCAAGCTGGGGTCTCACCAGCGACCGC 307

RESULT 6  
 BM501068 383 bp mRNA linear EST 14-FEB-2002  
 LOCUS PAC00000000933 Pioneer Af-1 array Zea mays cDNA, mRNA sequence.  
 DEFINITION BM501068  
 ACCESSION BM501068  
 VERSION BM501068.1 GI:18660825  
 KEYWORDS EST.  
 SOURCE zea mays.  
 ORGANISM zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 383)  
 Hunter,B.G., Beatty,M., Singletary,G., Hamaker,B., Larkins,B.A. and  
 Jung,R.  
 Maize opaque endosperm mutations create extensive changes in  
 patterns of gene expression  
 Unpublished (2002)  
 Contact: Jung R  
 Trait and Technology Development, Food and Feed Research  
 Pioneer Hi-Bred International, Inc.  
 7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA  
 Tel: 515 270 5934  
 Fax: 515 254 2619  
 Email: rudolf.jung@pioneer.com.  
 Location/Qualifiers  
 1..383  
 /organism="Zea mays"  
 /db\_xref="taxon:4577"  
 /clone\_11b="Pioneer Af-1 array"  
 /note="vector: pSPORT1; site\_1: SalI; site\_2: NotI"

BASE COUNT 78 a 112 c 103 g 77 t 11 others  
 ORIGIN

Query Match 6.3%; Score 85; DB 13; Length 383;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 GCTTAAGCAAGTAGCTTCATTACCCAGAGCGAGGACAGAGGAGCGCTTGCCATGGCG 72  
 DB 2 GCTTAAGCAAGTAGCTTCATTACCCAGAGCGAGGACAGAGGAGCGCTTGCCATGGCG 61

OY 73 TCTCCACCTGATGCAATGCTGCTGCGCTTCCTCCCTCCTGCTGCGCCACGCA 97  
 DB 62 TCTCCACCTGATGCAATGCTGCTGCGCTTCCTCCCTCCTGCTGCGCCACGCA 86

RESULT 7



BE362420 404 bp mRNA linear EST 20-JUL-2000  
 LOCUS DGL.85.C05.g1\_A002 Dark Grown 1 (DGL) Sorghum bicolor cDNA, mRNA  
 DEFINITION  
 ACCESION BE362420  
 VERSION BE362420.1 GI:9303977  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 L.H.  
 TITLE An EST database from Sorghum: dark-grown seedlings  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@pratt.uga.edu  
 Sequences have been trimmed to exclude polyA, vector and regions  
 below phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: PolyTm1x  
 High quality sequence start: 4  
 High quality sequence stop: 394  
 POLYA=Yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..404  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Dark Grown 1 (DGL)"  
 /note="Organ: 5-day-old dark-grown seedlings; Vector:  
 Lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was  
 made from poly-A RNA in the cloning vector lambda zap II.  
 Clones to be sequenced were prepared by mass excision."

BASE COUNT 87 a 127 c 122 g 68 t

ORIGIN

Query Match 5.6%; Score 76; DB 10; Length 404;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 817 CGGGGCGCTGTTCCACTCGGACGAGAGCTCTTCAACGGCGGTCGACAGCGCGTGGTG 876  
 |||||||  
 DB 115 CGGGGCGCTGTTCCACTCGGACGAGAGCTCTTCAACGGCGGTCGACAGCGCGTGGTG 174  
 |||||||

OY 877 AGGCAGTACAGCGCCA 892  
 |||||||

DB 175 AGGCAGTACAGCGCCA 190

RESULT 8  
 BE597204 436 bp mRNA linear EST 18-AUG-2000  
 LOCUS PTL.69.D10.g1\_A002 Pathogen Induced 1 (P11) Sorghum bicolor cDNA, mRNA  
 DEFINITION  
 ACCESION BE597204  
 VERSION BE597204.1 GI:9852179  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt  
 L.H.  
 TITLE An EST database from Sorghum: pathogen-induced plants

JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@pratt.uga.edu  
 Sequences have been trimmed to exclude polyA, vector and regions  
 below phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: PolyTm1x  
 High quality sequence start: 6  
 High quality sequence stop: 419  
 POLYA=Yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..436  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Pathogen induced 1 (P11)"  
 /note="Organ: Anthracnose-infected leaves from  
 two-week-old sorghum plants 48 hr after inoculation;  
 Vector: pBluescript II from lambda zap II; Site\_1: XhoI;  
 Site\_2: EcoRI; Two-week-old sorghum plants (Bry 623  
 cultivar) were infected with pathogen (isolate FM421 of  
 Colletotrichum graminicola, which is a sorghum isolate).  
 RNA was prepared from infected leaves harvested from 45  
 seedlings 48 hours after inoculation. Note: young  
 seedlings (2 weeks old) exhibit juvenile resistant  
 reaction, which is an incompatible interaction. As they  
 grow older (4 weeks or older), plants resume susceptibility  
 to anthracnose disease. The library was made from poly-A  
 RNA in the cloning vector lambda zap II. Clones to be  
 sequenced were prepared by mass excision. WARNING: While  
 most or all ESTs are expected to derive from the host  
 plant, no effort was made to eliminate ESTs deriving from  
 the pathogen."

BASE COUNT 97 a 135 c 127 g 77 t

ORIGIN

Query Match 5.6%; Score 76; DB 10; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 817 CGGGGCGCTGTTCCACTCGGACGAGAGCTCTTCAACGGCGGTCGACAGCGCGTGGTG 876  
 |||||||  
 DB 117 CGGGGCGCTGTTCCACTCGGACGAGAGCTCTTCAACGGCGGTCGACAGCGCGTGGTG 176  
 |||||||

OY 877 AGGCAGTACAGCGCCA 892  
 |||||||

DB 177 AGGCAGTACAGCGCCA 192

RESULT 9  
 BE362317 471 bp mRNA linear EST 20-JUL-2000  
 LOCUS DGL.85.C06.g1\_A002 Dark Grown 1 (DGL) Sorghum bicolor cDNA, mRNA  
 DEFINITION  
 ACCESION BE362317  
 VERSION BE362317.1 GI:9303874  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 L.H.  
 TITLE An EST database from Sorghum: dark-grown seedlings  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@prattuga.edu

Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.

Seq primer: PolYTmIX  
High quality sequence start: 40  
High quality sequence stop: 449  
POLYA-No.

## FEATURES

source location/Qualifiers

1..471  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Dark Grown 1 (DG1)"  
/note="Organ: 5-day-old dark-grown seedlings; Vector:  
lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was  
made from poly-A RNA in the cloning vector lambda Zap II.  
Clones to be sequenced were prepared by mass excision."  
BASE COUNT 106 a 147 c 134 g 84 t  
ORIGIN

Query Match 5.6%; Score 76; DB 10; Length 471;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 CGGGGCGCTGTCACCTCGACAGAGCTCTTCAACGCGGTCGACAGACGCGCTGGT 876  
|||||  
Db 156 CGGGGCGCTGTCACCTCGACAGAGAGCTCTTCAACGCGGTCGACAGACGCGCTGGT 215

QY 877 AGGAGTACAGCGCCA 892  
|||||  
Db 216 AGGAGTACAGCGCCA 231

## RESULT 10

BE597234 471 bp mRNA linear EST 18-AUG-2000  
LOCUS P11.69.G08.g1.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,  
DEFINITION mRNA sequence.

ACCESSION BE597234  
VERSION BE597234.1 GI:9852209  
KEYWORDS EST.

## SOURCE

ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.

## REFERENCE

1 (bases 1 to 471)  
Cordonier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt  
,L.H.

## AUTHORS

An EST database from Sorghum: pathogen-induced plants  
Unpublished (2000)

## JOURNAL

COMMENT Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@prattuga.edu

Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.

Seq primer: PolYTmIX  
High quality sequence start: 4  
High quality sequence stop: 443  
POLYA-Yes.

## FEATURES

source location/Qualifiers

1..471  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"

## BASE COUNT

121 a 128 c 131 g 91 t

Query Match 5.6%; Score 76; DB 10; Length 471;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 CGGGGCGCTGTCACCTCGACAGAGCTCTTCAACGCGGTCGACAGACGCGCTGGT 876  
|||||  
Db 83 CGGGGCGCTGTCACCTCGACAGAGAGCTCTTCAACGCGGTCGACAGACGCTGGT 142

QY 877 AGGAGTACAGCGCCA 892  
|||||  
Db 143 AGGAGTACAGCGCCA 158

## RESULT 11

BM318212 483 bp mRNA linear EST 03-JAN-2002  
LOCUS P11.79.G11.g9.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,  
DEFINITION mRNA sequence.

ACCESSION BM318212  
VERSION BM318212.1 GI:1805254  
KEYWORDS EST.

## SOURCE

ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.

## REFERENCE

1 (bases 1 to 483)  
Cordonier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt  
,L.H.

## AUTHORS

An EST database from Sorghum: pathogen-induced plants  
Unpublished (2000)

## JOURNAL

COMMENT Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@prattuga.edu

Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20. Three-prime sequences, which are obtained with PolYTmIX or  
T7 sequencing primer, are presented as the reverse complement.

Seq primer: PolYTmIX  
High quality sequence start: 3  
High quality sequence stop: 481  
POLYA-No.

## FEATURES

source location/Qualifiers

1..483  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Pathogen induced 1 (P11)"  
/note="Organ: Anthracnose-infected leaves from



Query Match 5.6%; Score 76; DB 10; Length 566;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 CGGGGCTGTTCACCTCGAGACGAGAGCTCTTCAACGCGGCGTCCGAGACGCGCTGCTG 876  
 |||||||  
 Db 78 CGGGGCTGTTCACCTCGAGACGAGAGCTCTTCAACGCGGCGTCCGAGACGCGCTGCTG 137  
 |||||||

QY 877 AGGCACTACAGCCCA 892  
 |||||||

Db 138 AGGCACTACAGCCCA 153

RESULT 14  
 BE597297 570 bp mRNA linear EST 18-AUG-2000-  
 LOCUS P11\_71.G06.g1.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,  
 DEFINITION mRNA sequence.  
 ACCESSION BE597297  
 VERSION BE597297.1 GI:9852370  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Sorghum.  
 1 (bases 1 to 570)  
 Cordonier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt  
 L.H.  
 An EST database from Sorghum: pathogen-induced plants  
 Unpublished (2000)  
 CONTACT: Cordonier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@prattuga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: Polymix  
 High quality sequence start: 37  
 High quality sequence stop: 547  
 POLYA-No.

FEATURES  
 source  
 1..570 Location/Qualifiers  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Pathogen induced 1 (P11)"  
 /note="Organ: Anthracnose-infected leaves from  
 two-week-old sorghum plants 48 hr after inoculation;  
 Vector: pBluescript II from lambda Zap II; Site.1: XhoI;  
 Site.2: EcoRI; Two-week-old sorghum plants (BRX 623  
 cultivar) were infected with pathogen (isolate PM421 of  
 Colletotrichum graminiicola, which is a sorghum isolate).  
 RNA was prepared from infected leaves harvested from 45  
 seedlings 48 hours after inoculation. Note: young  
 seedlings (2 weeks old) exhibit juvenile resistant  
 reaction, which is an incompatible interaction. As they  
 grow older (4 weeks or older), plants resume susceptibility  
 to anthracnose disease. The library was made from poly-A  
 RNA in the cloning vector lambda Zap II. Clones to be  
 sequenced were prepared by mass excision. WARNING: While  
 most or all ESTs are expected to derive from the host  
 plant, no effort was made to eliminate ESTs deriving from  
 the pathogen."

BASE COUNT 142 a 164 c 160 g 104 t  
 ORIGIN

Query Match 5.6%; Score 76; DB 10; Length 570;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 CGGGGCTGTTCACCTCGAGACGAGAGCTCTTCAACGCGGCGTCCGAGACGCGCTGCTG 876  
 |||||||  
 Db 154 CGGGGCTGTTCACCTCGAGACGAGAGCTCTTCAACGCGGCGTCCGAGACGCGCTGCTG 213  
 |||||||

QY 877 AGGCACTACAGCCCA 892  
 |||||||

Db 214 AGGCACTACAGCCCA 229

RESULT 15  
 BE367035 575 bp mRNA linear EST 20-JUL-2000-  
 LOCUS P11\_42.B10.g2.A002 Pathogen induced 1 (P11) sorghum bicolor cDNA,  
 DEFINITION mRNA sequence.  
 ACCESSION BE367035  
 VERSION BE367035.1 GI:9308592  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Sorghum.  
 1 (bases 1 to 575)  
 Cordonier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt  
 L.H.  
 An EST database from Sorghum: pathogen-induced plants  
 Unpublished (2000)  
 CONTACT: Cordonier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@prattuga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: Polymix  
 High quality sequence start: 14  
 High quality sequence stop: 573  
 POLYA-No.

FEATURES  
 source  
 1..575 Location/Qualifiers  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Pathogen induced 1 (P11)"  
 /note="Organ: Anthracnose-infected leaves from  
 two-week-old sorghum plants 48 hr after inoculation;  
 Vector: pBluescript II from lambda Zap II; Site.1: XhoI;  
 Site.2: EcoRI; Two-week-old sorghum plants (BRX 623  
 cultivar) were infected with pathogen (isolate PM421 of  
 Colletotrichum graminiicola, which is a sorghum isolate).  
 RNA was prepared from infected leaves harvested from 45  
 seedlings 48 hours after inoculation. Note: young  
 seedlings (2 weeks old) exhibit juvenile resistant  
 reaction, which is an incompatible interaction. As they  
 grow older (4 weeks or older), plants resume susceptibility  
 to anthracnose disease. The library was made from poly-A  
 RNA in the cloning vector lambda Zap II. Clones to be  
 sequenced were prepared by mass excision. WARNING: While  
 most or all ESTs are expected to derive from the host  
 plant, no effort was made to eliminate ESTs deriving from  
 the pathogen."

BASE COUNT 142 a 164 c 162 g 107 t  
 ORIGIN

Query Match 5.6%; Score 76; DB 10; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 CGGGGCTGTTCACCTCGAGACGAGAGCTCTTCAACGCGGCGTCCGAGACGCGCTGCTG 876  
 |||||||

Wed May 7 14:10:00 2003

us-10-047-825-3.rst

Page 9

Db 128 CGGGGCTGTTCACATCGACAGAGCTTTCACGGGGGTGCGAGACGCGCTGCTG 187  
QY 877 AGGCACTACAGGGCCA 892  
|||||  
Db 188 AGGCACTACAGGGCCA 203

Search completed: May 3, 2003, 11:27:31  
Job time : 1465 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: April 26, 2003, 12:20:37 ; Search time 77 Seconds

(Without alignments)  
541.655 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 1604

Sequence: 1 MASPTMQLCTVAIVSLSCVA.....VGVLTGTAGQIRNCRVNVS 313

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
A.GeneSeq\_101002:\*  
1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	982.5	61.3	327	22	AB99738
2	980	60.1	324	23	AB99735
3	965	60.2	318	21	AB16437
4	954.5	59.5	327	22	AB99741
5	950	59.2	323	21	AB16443
6	936.5	58.4	320	18	AA38217
7	926	57.7	314	22	AB99746
8	912.5	56.9	325	23	AB99743
9	901	56.2	326	22	AB99743
10	899.5	56.1	315	21	AB16426

11	885.5	55.2	320	22	AB99744	Oryza sativa perox
12	870.5	54.3	316	21	AA64911	Arabidopsis thalia
13	870.5	54.3	316	23	AB99736	Herb:cidally activ
14	867.5	54.1	316	21	AA64913	Arabidopsis thalia
15	866.5	54.0	308	21	AA64913	Arabidopsis thalia
16	866.5	54.0	309	21	AA64912	Arabidopsis thalia
17	863.5	53.8	308	21	AA64912	Arabidopsis thalia
18	863.5	53.8	309	21	AA64915	Arabidopsis thalia
19	846	52.7	266	21	AB16442	Pinus radiata pero
20	822	51.2	333	20	AA67892	Maize per root pr
21	818	51.0	311	23	AB97892	Lotium perenne per
22	813.5	50.7	321	23	AB91001	Herb:cidally activ
23	813	50.7	326	22	AB99742	Oryza sativa perox
24	808	50.4	331	23	AB99743	Herb:cidally activ
25	801	49.9	315	23	AB91000	Herb:cidally activ
26	801	49.9	338	21	AA62316	Arabidopsis thalia
27	801	49.9	338	21	AA64912	Arabidopsis thalia
28	801	49.9	338	23	AB91787	Herb:cidally activ
29	799	49.8	287	21	AB16435	Pinus radiata pero
30	778	48.5	346	21	AA60809	Arabidopsis thalia
31	774	48.3	337	21	AA630617	Herb:cidally activ
32	774	48.3	337	23	AB91786	Arabidopsis thalia
33	771.5	47.7	335	23	AB99737	Herb:cidally activ
34	765.5	47.1	358	21	AA64914	Herb:cidally activ
35	753.5	47.0	358	23	AB99737	Arabidopsis thalia
36	753.5	46.1	336	23	AB99737	Herb:cidally activ
37	739.5	46.1	336	21	AA64910	Arabidopsis thalia
38	739.5	46.1	336	23	AB99737	Herb:cidally activ
39	739.5	46.1	356	21	AA64910	Arabidopsis thalia
40	728.5	45.4	285	21	AA60810	Arabidopsis thalia
41	727.5	45.4	352	19	AA67138	Soybean seed coat
42	727.5	45.4	352	19	AA67138	Soybean seed coat
43	725	45.2	321	21	AA617363	Arabidopsis thalia
44	725	45.2	321	23	AB99749	Herb:cidally activ
45	725	45.2	328	21	AA617362	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1  
ID AB99738 standard; Protein; 327 AA.  
XX  
AC AB99738:  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Oryza sativa peroxidase s4235 protein SEQ ID NO:14.  
XX  
KW Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;  
modification; plant; bacterial infection; Magnaporthe grisea.  
XX  
OS Oryza sativa.  
XX  
XX WO200142475-A1.  
XX  
XX 14-JUN-2001.  
XX  
XX PD  
XX PF 08-DEC-2000; 2000WO-JP08728.  
XX PR 10-DEC-1999; 99JP-0352472.  
XX  
XX (NORU) JAPAN MIN AGRIC FORESTRY & FISHERIES.  
XX Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;  
XX Hiraga S;  
XX WPI: 2001-381695/40.  
XX DR N-PSDB: AAH44077.  
XX  
XX New set of rice peroxidase genes for analysis of peroxidase expression  
in rice under varying conditions and production of rice plants with

PT desired characteristics -  
 XX  
 PS Disclosure; Page 155-157; 258pp; Japanese.

CC The present invention describes a set of peroxidase genes found in  
 CC plants, especially rice, and their homologues, modified forms and  
 CC fragments, where the sequences of the peroxidase genes in the set are  
 CC given in AHH4071 to AHH4091. Also described are: (1) promoters for the  
 CC control of the gene set; (2) the preparation of cassette vectors using  
 CC the genes and promoters; (3) analysis of plant characteristics using the  
 CC peroxidase set by isolating RNA from the plant, binding the RNA to a  
 CC membrane, mixing with a labelled peroxidase gene set, incubating, and  
 CC detecting the label signal to show which genes in the set are expressed  
 CC in the sample plant; and (4) DNA microarrays for peroxidase gene  
 CC pattern of peroxidase gene expression in particular rice plants and  
 CC their component tissues and under different environmental conditions,  
 CC and modification of rice plants to provide desired specificities of  
 CC peroxidase gene expression to impart particular characteristics to the  
 CC plants such as response to bacterial infection by Magnaporthe grisea.  
 CC The present sequence represents a rice peroxidase encoded by a gene from  
 CC the gene set described above.

CC  
 XX  
 SQ Sequence 327 AA;

Query Match 61.3%; Score 982.5; DB 22; Length 327;  
 Best Local Similarity 61.5%; Pred. No. 3.5e-89;

Matches 192; Conservative 42; Mismatches 69; Indels 9; Gaps 3;

QY 11 VAVSLSCVANAQSLPFFYASCCPNLQSIYRAAMTQAVASEQMGASLRLFFHDCFYOG 70  
 DB 15 VAVVFLAEQSLSPGYNATCGVYIVRGMQAQVQESMGASILRLFFHCFYNG 74  
 QY 71 CGSITLD---AGGKTAGPNLNSVGFYDITIKRNEYAACPGVVCADITLAAADG 126  
 DB 75 CGASITLDDTANFTGKRNAGPNANSVGFYDITIKRNEYAACPGVVCADITLAAADG 134  
 QY 127 TNLGCPWVPLGRDSTASLANSNPPPTASLGLTISLFGHGLSPDMTALSGA 186  
 DB 135 VNLGCPWVPLGRDSTASLANSNPPPTASLGLTISLFGHGLSPDMTALSGA 194  
 QY 187 HTTGARCTTFRGRTYGGTDINAFSAF-LRQOTCPRSQGGNAPIDVOTPYRFTAYFT 245  
 DB 195 HTVGARCTTFRGRTYGGTDINAFSAF-LRQOTCPRSQGGNAPIDVOTPYRFTAYFT 254  
 QY 246 NLSRGLFHSQDEL-----NGSODALVROYASASLFPADPVAAIMTGVVLSTA 301  
 DB 255 DLSRGLFHSQDEL-----NGSODALVROYASASLFPADPVAAIMTGVVLSTA 314  
 QY 302 GOIRNRCRVVNS 313  
 DB 315 GEVIRNCRVNS 326

RESULT 2

ABB93359  
 ID ABB93359 standard; Protein: 324 AA.

XX  
 AC ABB93359;

DT 31-MAY-2002 (first entry)

XX  
 DE Herbicidally active polypeptide SEQ ID NO 2570.

XX  
 KW Herbicidal; plant; agriculture; herbicide.

XX  
 OS Arabidopsis thaliana.

XX  
 PN WO200210210-A2.

XX  
 PD 07-FEB-2002.

XX  
 PF 28-AUG-2001; 2001WO-EP09892.

XX  
 PR 28-AUG-2001; 2001WO-EP09892.

XX  
 PA (FARB ) BAYER AG.

XX  
 PI Tietjen K, Welder M;

XX  
 DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms -

PS Claim 5; SEQ ID NO 2570; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins  
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising  
 CC aligning and comparing nucleic acid or amino acid sequences from plant  
 CC with nucleic acid or amino acid sequences from non-plant organisms using  
 CC suitable search parameters, where plant sequences having an E-value  
 CC greater by a factor of 3 than the E-value of most similar non-plant  
 CC sequences are selected. The polypeptides or nucleic acids encoding them  
 CC are useful for identifying modulators. The identified modulators are  
 CC useful as herbicides.

CC  
 XX  
 SQ Sequence 324 AA;

Query Match 61.1%; Score 980; DB 23; Length 324;  
 Best Local Similarity 61.4%; Pred. No. 6.1e-89;

Matches 199; Conservative 36; Mismatches 77; Indels 12; Gaps 3;

QY 1 MASPTLMOCLYAVSL-----SCVAHOLSPPFYASCCPNLQSIYRAAMTQAVASEQRM 54  
 DB 1 MASPTLMOCLYAVSL-----SCVAHOLSPPFYASCCPNLQSIYRAAMTQAVASEQRM 60  
 QY 55 GASLRLFFHCFYOGCDGSLD---AGGKTAGPNLNSVGFYDITIKRNEYAAC 110  
 DB 61 GASLRLFFHCFYOGCDGSLD---AGGKTAGPNLNSVGFYDITIKRNEYAAC 120  
 QY 111 GVSCADITLAAARDGYNLGGPTWVPLGRDSTASLANSNPPPTASLGLTISLF 170  
 DB 121 GVSCADITLAAARDGYNLGGPTWVPLGRDSTASLANSNPPPTASLGLTISLF 180  
 QY 171 GNGGLSPDMTALSGARTTIGACTFRGRTYGGTDINAFSAF-LRQOTCPRSQGGN 228  
 DB 181 GNGGLSPDMTALSGARTTIGACTFRGRTYGGTDINAFSAF-LRQOTCPRSQGGN 240  
 QY 229 APIDVOTPYRFTAYFTNLSSRGLFHSQDELFGSODALVROYASASLFPADPVAAIM 288  
 DB 241 APIDVOTPYRFTAYFTNLSSRGLFHSQDELFGSODALVROYASASLFPADPVAAIM 300  
 QY 289 IRMGNVGLTGTAGOIRNRCRVVNS 312  
 DB 301 IRMGNVGLTGTAGOIRNRCRVVNS 324

RESULT 3

AAB16437  
 ID AAB16437 standard; Protein: 318 AA.

XX  
 AC AAB16437;

DT 24-OCT-2000 (first entry)

XX  
 DE Pinus radiata peroxidase protein sequence SEQ ID NO:389.

XX  
 KW Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;

XX  
 PN Pinus radiata; Monterey pine.

XX  
 PF WO200022099-A1.



XX 20-APR-2000.  
 XX 06-OCT-1999; 99WO-N200168.  
 XX 09-OCT-1998; 98US-0169789.  
 PR 14-JUL-1999; 99US-0143811.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 PI Bloksberg LN, Havukkala IU;  
 DR WPI; 2000-317962/27.  
 XX Novel polynucleotide encoding enzymes involved in lignin-biosynthetic  
 PT pathway useful for producing transgenic plants especially eucalyptus  
 PT and pine species having altered lignin content, composition and  
 PT structure  
 XX  
 XX Claim 18; Page 199; 213pp; English.  
 XX  
 XX The present invention describes isolated polynucleotides and proteins  
 CC encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),  
 CC coumarate 3-hydroxylase (C3H), phenolase (PNU), O-methyl transferase  
 CC (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase  
 CC (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),  
 CC coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),  
 CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,  
 CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,  
 CC coumarate CoA ligase, cytochrome P450 1X1A, diphenol oxidase, flavanol  
 CC glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase,  
 CC which are involved in the lignin biosynthetic pathway. The  
 CC polynucleotides can be used for modulating lignin content, lignin  
 CC composition and the structure of a plant, especially eucalyptus and pine  
 CC species, and for modifying the activity of an enzyme involved in lignin  
 CC biosynthetic pathway, and for producing a plant having altered lignin  
 CC content, composition and structure. They can be used for designing probes  
 CC and primers useful for detecting similar DNA and RNA sequences in any  
 CC organism and for PCR amplification. The lignin content can be efficiently  
 CC modified using the polynucleotides. AA67908 to AA68201 and AA616341 to  
 CC AA616449 represent polynucleotide and protein sequences used in the  
 CC exemplification of the present invention.  
 XX  
 XX Sequence 318 AA:  
 SQ  
 Query Match 60.2%; Score 965; DB 21; Length 318;  
 Best Local Similarity 60.1%; Pred. No. 1.9e-87;  
 Matches 191; Conservative 39; Mismatches 76; Indels 12; Gaps 2;  
 QY 7 MOCIVAVSLSC-----VAHQISPTTYASCPMLGSIYRAAMTQAVASQRMKASL 58  
 DB 1 MRLVLCIGMAVFAFHINAVNGSLSTTYASCPMLPBIYAVSVKQAAKERRKRGASL 60  
 QY LRLFFHDFVOCGDSIILDLG---GEKTAGPMLNSVRGFEYIDTIKRNVEAACPGVVS 114  
 DB 61 VRLHFDHDFVNCDSIILDLNATFTGKRTAGPRANASRGFEYIDSLIKQYVEAAGGVVS 120  
 QY 115 CADILAAAROGTNLGGPTWSVPLGRDSTTASASLANSNPPPTASGLTILSLFRGQ 174  
 DB 121 CADILTAARDSIVELGPTWVLMGRDSTTASLANNNIPSPASSLSTILSSFOAGS 180  
 QY 175 LSPRDMTALSGAHTTGACTTFFGRITGDTDIASFAALROOTCRSGSDGNLMAFDVQ 234  
 DB 181 LSTKDLVALSGAHTTGACAFPRIRIYNETINAAFAATSVKANCPSAGDSNLSPDLAV 240  
 QY 235 TVYRDETVATNLRLRSLFSDDELFGSGDIALVROYASASLNPADPVAMITMGAV 294  
 DB 241 TSTTPNKTKYSMLTKOKLHSDOGLFNGSDVQVAYVSNONSPTIDTAAVMKGIN 300  
 QY 295 GVLVTGAGQIRNRCRYVN 312  
 DB 301 SPLTGTNGQIRKRCRKN 318

RESULT 4  
 AAB99741  
 ID AAB99741 standard; Protein; 327 AA.  
 XX  
 AC AAB99741;  
 XX  
 DT 12-SEP-2001 (first entry)  
 DE Oryza sativa peroxidase r3025 protein SEQ ID NO:20.  
 XX  
 KM Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;  
 KM modification; plant; bacterial infection; Magnaporthe grisea.  
 XX  
 OS Oryza sativa.  
 XX  
 PN WO200142475-A1.  
 PD 14-JUN-2001.  
 XX  
 PF 08-DEC-2000; 2000WO-JP08728.  
 XX  
 PR 10-DEC-1999; 99JP-0352472.  
 XX  
 PA (NORO) JAPAN MIN AGRIC FORESTRY & FISHERIES.  
 XX  
 PI Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;  
 PI Hiraga S;  
 DR WPI; 2001-381695/40.  
 DR N-PSDB; AAH44080.  
 XX  
 PT New set of rice peroxidase genes for analysis of peroxidase expression  
 PT in rice under varying conditions and production of rice plants with  
 PT desired characteristics  
 XX  
 XX Disclosure; Page 174-176; 258pp; Japanese.  
 CC The present invention describes a set of peroxidase genes found in  
 CC plants, especially rice, and their homologues, modified forms and  
 CC fragments, where the sequences of the peroxidase genes in the set are  
 CC given in AAH44071 to AAH44091. Also described are: (1) promoters for the  
 CC control of the gene set; (2) the preparation of cassette vectors using  
 CC the genes and promoters; (3) analysis of plant characteristics using the  
 CC peroxidase set by isolating RNA from plants, binding the RNA to a  
 CC membrane, mixing with a labelled peroxidase gene set, incubating, and  
 CC detecting the label signal to show which genes in the set are expressed  
 CC in the sample plant; and (4) DNA microarrays for peroxidase gene  
 CC expression analysis. The set of genes are used for the analysis of the  
 CC pattern of peroxidase gene expression in particular rice plants and  
 CC their component tissues and under different environmental conditions,  
 CC and modification of rice plants to provide desired specificities of  
 CC peroxidase gene expression to impart particular characteristics to the  
 CC plants such as response to bacterial infection by Magnaporthe grisea.  
 CC The present sequence represents a rice peroxidase encoded by a gene from  
 CC the gene set described above.  
 XX  
 XX Sequence 327 AA:  
 SQ  
 Query Match 59.5%; Score 954.5; DB 22; Length 327;  
 Best Local Similarity 57.8%; Pred. No. 2.1e-86;  
 Matches 189; Conservative 46; Mismatches 77; Indels 15; Gaps 3;  
 QY 1 MASPT-----LMQCVAVSLSCVAHQISPTTYASCPMLGSIYRAAMTQAVAS 51  
 DB 1 MAOPTSARVRVRAALVAVVYVLAAGSMALDSPSTYSICPEVFAVAKRQMSALARE 60  
 QY 52 ORNGASLRLFFHDFVOCGDSIILD---AGEKTAGPMLNSVRGFEYIDTIKRNVEA 107  
 DB 61 KRIGASIVRLFFHDFVOCGDSIILDDTASFTGKRTANPNVGSVRGFEYIDAIKSAVER 120  
 QY 108 ACPGVVSCADILAAAROGTNLGGPTWSVPLGRDSTTASASLANSNPPPTASGLTIL 167

```

DB 121 ICPGVSCADILAIARSAVALIGSPMDVKVGRDRSRTSLSGANNPPPTSGIANLT 180
QY 168 SIFGRGIGSPDMTALSAHTITGOARCTTFRGRIYGTIDINAFALROOTCPRS--GGD 225
DB 181 SLFAAOALISQKQMVALLSGSTHTIGARCTNFRAHYNETNIDSGFAARRSGCGRNSGSD 240
QY 226 GNLAIPDVOTPVRFPTAFTNLRSRGLFHSQDELFGNGSODALVROYASASLIFNADIV 285
DB 241 NMLAPLDLOTPTVFENNRYKNLVYKGLLHSDQELFNGCATDIALVOSYISSGSTFFRADIV 300
QY 286 AAMINGMNVGLTGTAGQIRRCRVN 312
DB 301 TGMKMGDTITPLTSGNGEIRKRCRIN 327

RESULT 5
AAB16443
ID AAB16443 standard; Protein: 323 AA.
XX
AC AAB16443;
XX
DT 24-OCT-2000 (first entry)
XX
DE Pinus radiata peroxidase protein sequence SEQ ID NO:395.
XX
KW plant; lignin biosynthetic pathway; Eucalyptus grandis;
KW Pinus radiata; Monterey pine.
XX
OS Pinus radiata.
XX
PN MO200022099-AL.
XX
PD 20-APR-2000.
XX
PF 06-OCT-1999; 99WO-NZ00168.
XX
PR 09-OCT-1998; 98US-0169789.
PR 14-JUL-1999; 99US-0143811.
XX
PA (GENE-) GENESTS RES 6 DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN, Havukkala IJ;
XX
DR WPI; 2000-317962/27.
XX
PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic
PT pathway useful for producing transgenic plants especially eucalyptus
PT and pine species having altered lignin content, composition and
PT structure
XX
PS Claim 18; Page 202-203; 213pp; English.
XX
XX
XX The present invention describes isolated polynucleotides and proteins
XX encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),
XX coumarate 3-hydroxylase (C3H), phenolase (PMT), O-methyl transferase
XX (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase
XX (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),
XX coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),
XX laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,
XX caffeic acid methyl transferase, caffeoyl CoA methyl transferase,
XX coumarate CoA ligase, cytochrome P450 1X1A, diphenol oxidase, flavanol
XX glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase,
XX which are involved in the lignin biosynthetic pathway. The
XX polynucleotides can be used for modulating lignin content, lignin
XX composition and the structure of a plant, especially eucalyptus and pine
XX species, and for modifying the activity of an enzyme involved in lignin
XX biosynthetic pathway, and for producing a plant having altered lignin
XX content, composition and structure. They can be used for designing probes
XX and primers useful for detecting similar DNA and RNA sequences in any
XX organism and for PCR amplification. The lignin content can be efficiently
XX modified using the polynucleotides. AAB67908 to AAB68201 and AAB16341 to

```

```

CC AAB16449 represent polynucleotide and protein sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 323 AA:
QY 9 CLVAASLLSC---VAHAQLSPTFYAASCPNLGSIIVRAAMTQAAVASBQRMGASLLRLFEH 64
DB 10 CTAIVWFVTCSTIDYVNGQLSTFYDKCPITALSVYAAVQAQAAVANKRRGASLLRLHFF 69
QY 65 DCFVQGCDSILLDAG---GKTAGPNLNSVRGEVIDITKRNVEACPGVSCADIIA 120
DB 70 DCFVNGCDSVLLDSSITTGERTANPNANSAARGFDVIDITKSNVEKACGVSCADIIA 129
QY 121 LAARDGNTLLGPTWSPVPLGRDRSTTASASLANSNPPPTSLGLTSLFGROGLSPDM 180
DB 130 IAAKDSVVELGSPSWTVMAGRDRSTTASASLANSNPPPTSLGLTSLFGROGLSAKEM 189
QY 181 TALSGAHTIGARCTTFRGRIYGTIDINAFALROOTCPRS--GGDGNLAPDVOTPV 238
DB 190 VALSGHTIGAOCKNFRAHYNETNIDSAVATSLRSKCPSTGSGDSNLEPLDYMTPTV 249
QY 239 FDMAYFTNLRSRGLFHSQDELFGNGSODALVROYASASLIFNADIVAAIMIRGNVGLT 298
DB 250 FDKNYTSDLSQKGLLHSDQELFNGSDVQVTTASNQNTFESDFAAAMKMNINPLT 309
QY 299 GTAGQIRRCRVN 312
DB 310 GTSGQIPKNCRKPN 323

RESULT 6
AAM38217
ID AAM38217 standard; Protein: 320 AA.
XX
AC AAM38217;
XX
DT 27-APR-1998 (first entry)
XX
DE Stylosanthes humilis Shpx6 peroxidase.
XX
KW Peroxidase; Shpx6; transgenic plant; fungus resistance;
KW disease protection; Phytophthora parasitica;
KW Leptosphaeria maculans; Sclerotinia sclerotiorum.
XX
OS Stylosanthes humilis straln Paterson.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT /label= Sig-peptide
XX
XX WO9741237-A1.
XX
XX 06-NOV-1997.
XX
XX 29-APR-1997; 97WO-AU00253.
XX
XX 29-APR-1996; 96AU-0009532.
XX
XX (EETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX (GRAI-) GRAINS RES 6 DEV CORP.
XX
XX Goulter KC, Kazan K, Manners JM;
XX
XX WPI; 1997-549739/50.
XX
XX N-PSDB; AAT95782.
XX
XX Plant cell transformed with Stylosanthes humilis Shpx6 peroxidase
XX useful to confer on plants resistance to fungi, e.g. Phytophthora
XX parasitica, Leptosphaeria maculans or Sclerotinia sclerotiorum

```

XX Claim 4; Page 15-16; 30pp; English.  
PS  
XX This protein comprises the Shp6 peroxidase of Stylosanthes  
CC humilis. A claimed method of engineering a plant to fungal  
CC resistance comprises introducing into cells of the plant a DNA  
CC construct comprising a promoter that is constitutively operative  
CC in the plant cell (preferably the cauliflower mosaic virus 35S  
CC promoter) and a Shp6 DNA sequence (see AAT95782), or a hybridising  
CC sequence or fragment that encodes an enzyme with peroxidase  
CC activity. Also claimed are plant cells harboring the DNA  
CC construct, a plant comprising such cells, and material from such  
CC a plant, especially seed, pollen, a stem segment or a cutting.  
CC Resistance to fungi, especially Phytophthora parasitica,  
CC Leptosphaeria maculans or Sclerotinia sclerotiorum, can be  
CC conferred on cereal, legume, oilseed, sugar or fibre plants,  
CC particularly maize, banana, peanut, field pea, sunflower, tomato,  
CC canola, tobacco, wheat, barley, oat, potato, soybean, cotton,  
CC carnation and sorghum.  
XX  
SQ Sequence 320 AA;  
Query Match 58.4%; Score 936.5; DB 18; Length 320;  
Best Local Similarity 59.9%; Pred. No. 1.3e-84;  
Matches 187; Conservative 41; Mismatches 75; Indels 9; Gaps 4;  
QY 9 CLV--AVSLSCVAHAQISPTFYASSCPNIQSIYRAMTQAVASEQMGSLIRFPHC 66  
DB 10 CLILVMSLIQ-LGSGQLSNFVATCPNALSTIRSVNSAVSKAEAMGSLIRLPHDC 68  
QY 67 FVQGGDSIILD---AGGECTAGPNLNSVGRFEVIDTIKRNVEAACPGVSCADILALA 122  
DB 69 FVQGGDSAVILDDPSNFGEXTARPANNSINGFEVIDTIKSYVESLCPGVSCADILAVA 128  
QY 123 ARDGTNLGGPTWSPVLRGRDSTTASASLANSNPPTASLIGLISLFGQGLSPRDMTA 182  
DB 129 AADSVVALGSPWTVLGRDSTTASASLANSDLAAPTLLDLSLGLISLFGQGLSTSEMA 188  
QY 183 LSGAHTIGQARCTTFRGRITGDTIDINAFALRQOTCPRS--GGDGNLAPIDVQTPVRF 240  
DB 189 LSGHTIGQARCTTFRGRITGDTIDINAFALRQOTCPRS--GGDGNLAPIDVQTPVRF 248  
QY 241 TAYFTNLISRGLEFHSODELFNGSGDALVROYASASLFPNADVAMRMNGVGLTGT 300  
DB 249 NCYVNLILVKGLFHSIDOLFNGSGDVSQVNGVANSFSCDFGNMIMKGNISPLTGS 308  
QY 301 AGQIRNCRVYN 312  
DB 309 SGQIRNCRKTN 320  
RESULT 7  
AAB99746  
ID AAB99746 standard; Protein: 314 AA.  
XX  
AC AAB99746;  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Oryza sativa peroxidase r2576 protein SEQ ID NO:30.  
XX  
KW Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;  
XX modification; plant; bacterial infection; Magnaporthe grisea.  
OS Oryza sativa.  
XX  
XX WO200142475-A1.  
PN  
PD 14-JUN-2001.  
XX  
PF 08-DEC-2000; 2000WO-JP08728.  
XX  
PR 10-DEC-1999; 99JP-0352472.

XX  
PA (NORQ ) JAPAN MIN AGRIC FORESTRY & FISHERIES.  
XX  
PI Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;  
XX Hirata S;  
XX  
DR WPI: 2001-381695/40.  
DR N-PSDB: AAH44085.  
XX  
PT New set of rice peroxidase genes for analysis of peroxidase expression  
PT in rice under varying conditions and production of rice plants with  
PT desired characteristics  
PS  
XX Disclosure; Page 205-207; 258pp; Japanese.  
XX  
XX The present invention describes a set of peroxidase genes found in  
CC plants, especially rice, and their homologues, modified forms and  
CC fragments, where the sequences of the peroxidase genes in the set are  
CC given in AAH44071 to AAH44091. Also described are: (1) promoters for the  
CC control of the gene set; (2) the preparation of cassette vectors using the  
CC the genes and promoters; (3) analysis of plant characteristics using the  
CC peroxidase set by isolating RNA from the plant, binding the RNA to a  
CC membrane, mixing with a labelled peroxidase gene set, incubating, and  
CC detecting the label signal to show which genes in the set are expressed  
CC in the sample plant; and (4) DNA microarrays for peroxidase gene  
CC expression analysis. The set of genes are used for the analysis of the  
CC pattern of peroxidase gene expression in particular rice plants and  
CC their component tissues and under different environmental conditions,  
CC and modification of rice plants to provide desired specificities of the  
CC peroxidase gene expression to impart particular characteristics to the  
CC plants such as response to bacterial infection by Magnaporthe grisea.  
CC The present sequence represents a rice peroxidase encoded by a gene from  
CC the gene set described above.  
SQ Sequence 314 AA;  
Query Match 57.7%; Score 926; DB 22; Length 314;  
Best Local Similarity 59.1%; Pred. No. 1.4e-83;  
Matches 185; Conservative 42; Mismatches 82; Indels 4; Gaps 3;  
QY 2 ASPTLMQCLVAVSLSCVAHAQISPTFYASSCPNIQSIYRAMTQAVASEQMGSLIRL 61  
DB 4 ASSVSLMLLVAAAAS-AASQSLATFTYDPSCPALSTISAVTAANSPRRGASLVR 62  
QY 62 FPHDCFVQGGDSIILDAGGECTAGPNLNSVGRFEVIDTIKRNVEAACPGVSCADILAL 121  
DB 63 FPHDCFVQGGDSIILDAGGECTAGPNLNSVGRFEVIDTIKRNVEAACPGVSCADILAL 121  
QY 122 AARDGTNLGGPTWSPVLRGRDSTTASASLANSNPPTASLIGLISLFGQGLSPRDMT 181  
DB 122 AARDGTNLGGPTWSPVLRGRDSTTASASLANSNPPTASLIGLISLFGQGLSPRDMT 181  
QY 182 ALSGAHTIGQARCTTFRGRITGDTIDINAFALRQOTCPRS--GGDGNLAPIDVQTPVRF 239  
DB 182 ALSGAHTIGQARCTTFRGRITGDTIDINAFALRQOTCPRS--GGDGNLAPIDVQTPVRF 241  
QY 240 DTAFTNLISRGLEFHSODELFNGSGDALVROYASASLFPNADVAMRMNGVGLTGT 299  
DB 242 DTAFTNLISRGLEFHSODELFNGSGDVSQVNGVANSFSCDFGNMIMKGNISPLTGS 301  
QY 300 TAGQIRNCRVYN 312  
DB 302 TCGQIRNCRKTN 314  
RESULT 8  
AAB93867  
ID AAB93867 standard; Protein: 325 AA.  
XX  
AC AAB93867;  
XX  
DT 31-MAY-2002 (first entry)  
XX

DE Herbicidally active polypeptide SEQ ID NO 3078.  
XX  
XX  
KM Herbicidal; plant; agriculture; herbicide.  
OS Arabidopsis thaliana.  
XX  
PN WO200210210-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 28-AUG-2001; 2001MO-EP09892.  
XX  
PR 28-AUG-2001; 2001MO-EP09892.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Tjetjen K, Weidler M;  
XX  
DR WPI: 2002-269010/31.  
XX  
XX Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms -  
XX  
XX  
PS Claim 5; SEQ ID NO 3078; 261pp + Sequence Listing; English.  
XX  
XX The invention relates to identifying target proteins  
CC (ABB9790-ABB94016) for herbicidally active compounds, comprising  
CC aligning and comparing nucleic acid or amino acid sequences from plant  
CC with nucleic acid or amino acid sequences from non-plant organisms using  
CC suitable search parameters, where plant sequences having an E-value  
CC greater by a factor of 3 than the E-value of most similar non-plant  
CC sequences are selected. The polypeptides or nucleic acids encoding them  
CC are useful for identifying modulators. The identified modulators are  
CC useful as herbicides.  
XX  
XX  
SQ Sequence 325 AA;  
  
Query Match 56.9%; Score 912.5; DB 23; Length 325;  
Best Local Similarity 60.3%; Pred. No. 3.2e-82;  
Matches 187; Conservative 35; Mismatches 81; Indels 7; Gaps 3;  
  
QY 10 LVAVALSCVAHAQSPFFYASCPNLSQSIYRAAMTQAVASEQMGASLILFFHDFVQ 69  
DB 16 LFIYMLSQAOQARLTFYDSCSLPVRVYQREVAERILASLILFFHDFCVN 75  
QY 70 GCDGSILLDAC---GEKTAGPNLNSVRGEVIDTIKRNVEACPGVYSCADIIALARD 125  
DB 76 GCDASILLDDRSEFLGEKTAGPNNNSVRGEVIDAIKRSVERLCPGVSCADIIALARD 135  
QY 126 GYNLLGGPTWVPLGRDSTTASASLANSNP-PPPTASLGLTILSLFGQGLSPRDMTALS 184  
DB 136 SYLLMGSGWVYKLRDSTTASSTANSGLVLPPTSTLNLINIFRANGSLPRDMVALS 195  
QY 185 GAHTTGOARCTFFRGRIYGDIDINAFALROQTCPRS--GGDGNLADIVQTPVEDTA 242  
DB 196 GAHTTGOARCTFFRGRIYVSTNIDLSFALSRSCPAATGSDNNAAILDLTEPEKFDGS 255  
QY 243 YFTNLISRGLEFHSDOELFNGSQALVROYASASLSLNADPVMAMIMGVAVLTGAG 302  
DB 256 YFMOVLVNHGRLTSDOVLFNGSGSDTSIVYSRVQAEYRDEVAAMIMGDISPLTGSNG 315  
QY 303 QIRNRCRYVN 312  
DB 316 QIRRSCTRRPN 325

RESULT 9  
AAB99743  
ID AAB99743 standard: Protein; 326 AA.  
XX  
AC AAB99743;

XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Oryza sativa peroxidase s10927 protein SEQ ID NO:24.  
XX  
XX Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;  
KM modification; plant; bacterial infection; Magnaporthe grisea.  
XX  
OS Oryza sativa.  
XX  
PN WO200142475-A1.  
XX  
PD 14-JUN-2001.  
XX  
PF 08-DEC-2000; 2000WO-JP08728.  
XX  
PR 10-DEC-1999; 99JP-0352472.  
XX  
PA (NORQ ) JAPAN MIN AGRIC FORESTRY & FISHERIES.  
XX  
PI Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;  
PI Hiraga S;  
XX  
DR WPI: 2001-381695/40.  
XX  
DR N-PSDB; AAH44082.  
XX  
XX New set of rice peroxidase genes for analysis of peroxidase expression  
PT in rice under varying conditions and production of rice plants with  
PT desired characteristics -  
XX  
XX  
PS Disclosure; Page 186-188; 258pp; Japanese.  
XX  
XX  
XX The present invention describes a set of peroxidase genes found in  
CC plants, especially rice, and their homologues, modified forms and  
CC fragments where the sequences of the peroxidase genes in the set are  
CC given in AAH44071 to AAH44091. Also described are: (1) promoters for the  
CC control of the gene set; (2) the preparation of cassette vectors using the  
CC the genes and promoters; (3) analysis of plant characteristics using the  
CC peroxidase set by isolating RNA from the plant, binding the RNA to a  
CC membrane, mixing with a labelled peroxidase gene set, incubating, and  
CC detecting the label signal to show which genes in the set are expressed  
CC in the sample plant; and (4) DNA microarrays for peroxidase gene  
CC expression analysis. The set of genes are used for the analysis of the  
CC pattern of peroxidase gene expression in particular rice plants and  
CC their component tissues and under different environmental conditions,  
CC and modification of rice plants to provide desired specificities of  
CC the peroxidase gene expression to impart particular characteristics to the  
CC plants such as response to bacterial infection by Magnaporthe grisea.  
CC The present sequence represents a rice peroxidase encoded by a gene from  
CC the gene set described above.  
XX  
XX  
SQ Sequence 326 AA;  
  
Query Match 56.2%; Score 901; DB 22; Length 326;  
Best Local Similarity 54.1%; Pred. No. 4.5e-81;  
Matches 170; Conservative 61; Mismatches 77; Indels 6; Gaps 3;  
  
QY 6 LMOCLVAVSL-LSCVAHAQSPFFYASCPNLSQSIYRAAMTQAVASEQMGASLILFFH 64  
DB 13 LILLLAVALLAARAARALSPGFYSASCPYVGVYQVNSQAVMNDTRGAVLFLFH 72  
QY 65 DCFVGGCDGSILLD---AGGEKTAGPN-LNSVRGEVIDTIKRNVEACPGVYSCADII 119  
DB 73 DCFVGGCDASVLLDDTPAAAGKEKGVGNNAVSTTVEDVITIAQYEAQVAPATVSCADVL 132  
QY 120 ALAARDGYNLLGGPTWVPLGRDSTTASASLANSNP-PPPTASLGLTILSLFGQGLSPRD 179  
DB 133 ALAARDGYNLLGGPTWVPLGRDSTTASASLANSNP-PPPTASLGLTILSLFGQGLSPRD 192  
QY 180 MTALSGAHTTGOARCTFFRGRIYGDIDINAFALROQTCPRS--GGDGNLADIVQTPVE 239  
DB 193 LAAISGAHTVGRASCNVFRTYVCDANVSPAFASHQSCPASGDAALAPLDSLTIPDAF 252

QY 240 DTAFTNTLSRRGLFHSDOELFNGSODALVROYSASASLFLNADPVAAMIRKGNVLTG 299  
 DB 253 DNGYRNIVAGAGLHSDQELFNGPVDVYGLYSSNAAFSSDPAASWIRLNGPLTG 312  
 OY 300 TAGOIRRCRCRYVNS 313  
 DB 313 STGEVRLNCRKRVNS 326

## RESULT 10

AAB16426  
 ID AAB16426 standard; Protein; 315 AA.

AC AAB16426;

DT 24-OCT-2000 (first entry)

DE Eucalyptus grandis peroxidase protein sequence SEQ ID NO:378.

KW Plant; Lignin; Lignin biosynthetic pathway; Eucalyptus grandis;  
 KW Pinus radiata; Monterey pine.

OS Eucalyptus grandis.

PN WO200022099-A1.

PD 20-APR-2000.

PF 06-OCT-1999; 99WO-NZ00168.

PR 09-OCT-1998; 98US-0169789.

PR 14-JUL-1999; 99US-0143811.

PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

PI Bloksberg LN, Havukkala IJ;

DR WPI; 2000-317962/27.

PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic  
 PT pathway useful for producing transgenic plants especially eucalyptus  
 PT and pine species having altered lignin content, composition and  
 PT structure

PS Claim 18; Page 193; 213pp; English.

CC The present invention describes isolated polynucleotides and proteins  
 CC encoding and representing the enzymes cinnamate 4-hydroxylase (CAH),  
 CC coumarate 3-hydroxylase (C3H), phenolase (PML), O-methyl transferase  
 CC (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase  
 CC (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),  
 CC coniferol glucosyl transferase (CCR), coniferin beta-glucosidase (CBG),  
 CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,  
 CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,  
 CC coumarate CoA ligase, cytochrome P450 Lx1A, diphenol oxidase, flavanol  
 CC glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase,  
 CC which are involved in the lignin biosynthetic pathway. The  
 CC polynucleotides can be used for modulating lignin content, lignin  
 CC composition and the structure of a plant, especially eucalyptus and pine  
 CC species, and for modifying the activity of an enzyme involved in lignin  
 CC biosynthetic pathway, and for producing a plant having altered lignin  
 CC content, composition and structure. They can be used for designing probes  
 CC and primers useful for detecting similar DNA and RNA sequences in any  
 CC organism and for PCR amplification. The lignin content can be efficiently  
 CC modified using the polynucleotides. AAB67908 to AAB68201 and AAB16341 to  
 CC AAB16449 represent polynucleotide and protein sequences used in the  
 CC exemplification of the present invention.

XX Sequence 315 AA;

Query Match 56.1%; Score 899.5; DB 21; Length 315;  
 Best Local Similarity 57.0%; Pred. No. 6,1e-81;

Matches 180; Conservative 46; Mismatches 85; Indels 5; Gaps 2;  
 OY 1 MASPTIMQCELVAAVSLSCVAHAQDLPFPASSCPNLQSIIVRAAMQAVASQCBGASILR 60  
 DB 1 MVSFSVVVVLATISVIT-7ARCKLSPSHQSTCPALSIYRAGVAKAIKNTFRGASILR 59  
 OY 61 LFFHDCFYOGCDGSIILIDAG---GEKTAGPMLNSVRGEVYIDTIRKNEACPGVSCA 116  
 DB 60 LFFHDCFYNGCDASILLDDPSPFEVGEKTAAPNNNSVREGVEYIDRIKASLEKCGVSCA 119  
 OY 117 DILAAARDGTNLLGPTWGSVPLGRDSTTASASLANSNPPPTSLGTLSLGRGSL 176  
 DB 120 DIVAALANDSVYHLGSPSWVSLGRKDSITASRLANISIPPSINLSALTSPAAGLS 179  
 OY 177 PRDMTALSQAHTIGQARCTFRGRIYGDPTDINAPRALROQTCRPSGSGGULAPIDVQTP 236  
 DB 180 VKNMVALSGSHFTIGLARTCFRRRIYNSNIDTSPFAHKIQCPIGINDSVLQRLDIQTP 239  
 OY 237 VRFDTAFTNTLSRRGLFHSDOELFNGSODALVROYSASASLFLNADPVAAMIRKGNV 296  
 DB 240 TFFDMLYVHNLLQKGLHSDQELFNGSSVDSLVRKYACDCKFRPRAKAMIKSEIKP 299  
 OY 297 LTGTAQOIRRCRCRYVNS 312  
 DB 300 PKGSGOIRKRCRVNS 315

## RESULT 11

AAB99744  
 ID AAB99744 standard; Protein; 320 AA.

AC AAB99744;

DT 12-SEP-2001 (first entry)

DE Oryza sativa peroxidase p14493 protein SEQ ID NO:26.

KW Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;  
 KW modification; plant; bacterial infection; Magnaporthe grisea.

OS Oryza sativa.

PN WO200142475-A1.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-JP08728.

PR 10-DEC-1999; 99JP-0352472.

PA (NORO) JAPAN MIN AGRIC FORESTRY & FISHERIES.

PI Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;  
 PI Hiraga S;

DR WPI; 2001-381695/40.  
 DR N-PSDB; AAB44083.

PT New set of rice peroxidase genes for analysis of peroxidase expression  
 PT in rice under varying conditions and production of rice plants with  
 PT desired characteristics  
 PS Disclosure; Page 192-194; 258pp; Japanese.

CC The present invention describes a set of peroxidase genes found in  
 CC plants, especially rice, and their homologues, modified forms and  
 CC fragments, where the sequences of the peroxidase genes in the set are  
 CC given in AAB44071 to AAB44091. Also described are: (1) promoters for the  
 CC control of the gene set; (2) the preparation of cassette vectors using  
 CC the genes and promoters; (3) analysis of plant characteristics using the  
 CC peroxidase set by isolating RNA from the plant, binding the RNA to a  
 CC membrane, mixing with a labelled peroxidase gene set, incubating, and  
 CC detecting the label signal to show which genes in the set are expressed



PR	14-OCT-1999	99US-0159330.
PR	14-OCT-1999	99US-0159331.
PR	14-OCT-1999	99US-0159637.
PR	14-OCT-1999	99US-0159638.
PR	18-OCT-1999	99US-0159584.
PR	21-OCT-1999	99US-0160741.
PR	21-OCT-1999	99US-0160747.
PR	21-OCT-1999	99US-0160758.
PR	21-OCT-1999	99US-0160770.
PR	21-OCT-1999	99US-0160814.
PR	21-OCT-1999	99US-0160815.
PR	22-OCT-1999	99US-0160980.
PR	22-OCT-1999	99US-0160981.
PR	22-OCT-1999	99US-0160989.
PR	22-OCT-1999	99US-0161404.
PR	25-OCT-1999	99US-0161405.
PR	25-OCT-1999	99US-0161406.
PR	26-OCT-1999	99US-0161359.
PR	26-OCT-1999	99US-0161360.
PR	28-OCT-1999	99US-0161361.
PR	28-OCT-1999	99US-0161920.
PR	28-OCT-1999	99US-0161992.
PR	28-OCT-1999	99US-0161993.
PR	29-OCT-1999	99US-0162142.

Query Match:	54.3%;	Score 870.5;	DB 21;	Length 316;
Best Local Similarity	57.1%;	Pred. No. 4,7e-78;		
Matches 177;	Conservative 44;	Mismatches 82;	Indels 7;	Gaps 3;

  

QY	10	LVAVSLISCVAHNOISPTFYAASGCGNLSQIYRAMQTVASGEOMGSLILRFPHDCFYQ	69
DB	7	LAMAMMASGSEKQLNDRDPTKESCSLELVAKRYKRAAREPMGASLILFFHDCFYQ	66
QY	70	GCDGSILLDAG---GEKTAGPNLNSVGEFVIDTKIRNVEAACPGVSCAJTILAAKD	125
DB	67	GCQGSILLDPTPSFLQEKTSQGSNNNSVGFVEYDIKIRFEKKMCPGIVSCADTLATTD	126
QY	126	GTNLLGPTVSYVLGHRDSTASTASLANSNP-PPPTASLTGLISLFGQGLSPRDMTALS	184
DB	127	SVLLLGPGSVKVLGRDSTTANFPAANSQVIPPITTLNLIWRFAQGLSTRDWALIS	186
QY	185	GAHTIGGARTPTFRGRIVGDTIDINAFPAALROQTCPSRGGDGN--LAPIDVQTPVREFTA	242
DB	187	GAHTIGGAQCVTFTRNRIYANSLIIDSPALSKRRNCPTASSGGCDKKKNLDVRSDDFEDG	246
QY	243	YFTNLSLRGLPHSDDELFGSGQALYROYASASLSLFNADPFAAMTRMGNVGLTGTG	302
DB	247	FYKQLSKRKLISDQVLFENNGPTSLVIAYSHNLMAFYRDFAPAMIKMGDISPLTGSNG	306
QY	303	QIRRMCRVYN 312	
DB	307	QIRNCRFRPN 316	

  

RESULT 13	
ABB93866	
ID	ABB93866 standard; Protein; 316 AA.
XX	ABB93866;
AC	
XX	31-MAY-2002 (first entry)
XX	
DE	Herbicidally active polypeptide SEQ ID NO 3077.
XX	
XX	Herbicidal; plant; agriculture; herbicide.
XX	
OS	Arabidopsis thaliana.
XX	
PN	MO200210210-A2.
XX	
PD	07-FEB-2002.
XX	
FE	28-AUG-2001.2001MO-EP09892.





PR	01-JUL-1999	99US-0142154
PR	02-JUL-1999	99US-0142055
PR	05-JUL-1999	99US-0142800
PR	06-JUL-1999	99US-0142820
PR	09-JUL-1999	99US-0142820
PR	12-JUL-1999	99US-0142820
PR	13-JUL-1999	99US-0143342
PR	14-JUL-1999	99US-0143624
PR	15-JUL-1999	99US-0144005
PR	16-JUL-1999	99US-0144086
PR	19-JUL-1999	99US-0144335
PR	19-JUL-1999	99US-0144332
PR	19-JUL-1999	99US-0144332
PR	19-JUL-1999	99US-0144334
PR	19-JUL-1999	99US-0144352
PR	20-JUL-1999	99US-0144352
PR	20-JUL-1999	99US-0144684
PR	21-JUL-1999	99US-0144684
PR	21-JUL-1999	99US-0144511
PR	21-JUL-1999	99US-0145088
PR	22-JUL-1999	99US-0145088
PR	22-JUL-1999	99US-0145087
PR	22-JUL-1999	99US-0145192
PR	22-JUL-1999	99US-0145192
PR	23-JUL-1999	99US-0145145
PR	23-JUL-1999	99US-0145218
PR	23-JUL-1999	99US-0145324
PR	26-JUL-1999	99US-0145376
PR	27-JUL-1999	99US-0145513
PR	27-JUL-1999	99US-0145518
PR	27-JUL-1999	99US-0145519
PR	28-JUL-1999	99US-0145519
PR	02-AUG-1999	99US-0146386
PR	02-AUG-1999	99US-0146388
PR	02-AUG-1999	99US-0146388
PR	03-AUG-1999	99US-0147004
PR	04-AUG-1999	99US-0147004
PR	05-AUG-1999	99US-0147192
PR	05-AUG-1999	99US-0147303
PR	06-AUG-1999	99US-0147303
PR	06-AUG-1999	99US-0147463
PR	09-AUG-1999	99US-0147935
PR	10-AUG-1999	99US-0148171
PR	11-AUG-1999	99US-0148319
PR	12-AUG-1999	99US-0148341
PR	13-AUG-1999	99US-0148565
PR	16-AUG-1999	99US-0148684
PR	17-AUG-1999	99US-0149175
PR	18-AUG-1999	99US-0149426
PR	20-AUG-1999	99US-0149722
PR	20-AUG-1999	99US-0149723
PR	20-AUG-1999	99US-0149829
PR	23-AUG-1999	99US-0149802
PR	23-AUG-1999	99US-0149930
PR	25-AUG-1999	99US-0150566
PR	26-AUG-1999	99US-0150884
PR	27-AUG-1999	99US-0151065
PR	27-AUG-1999	99US-0151065
PR	27-AUG-1999	99US-0151083
PR	30-AUG-1999	99US-0151300
PR	31-AUG-1999	99US-0151330
PR	01-SEP-1999	99US-0151530
PR	07-SEP-1999	99US-0152263
PR	10-SEP-1999	99US-0153070
PR	13-SEP-1999	99US-0153758
PR	15-SEP-1999	99US-0154018
PR	16-SEP-1999	99US-0154018

PR	20-SEP-1999:	99US-0154779.
PR	22-SEP-1999:	99US-0155139.
PR	23-SEP-1999:	99US-0155486.
PR	24-SEP-1999:	99US-0155659.
PR	28-SEP-1999:	99US-0156458.
PR	29-SEP-1999:	99US-0156596.
PR	04-OCT-1999:	99US-0157117.
PR	05-OCT-1999:	99US-0157753.
PR	06-OCT-1999:	99US-0157865.
PR	07-OCT-1999:	99US-0158029.
PR	08-OCT-1999:	99US-0158232.
PR	12-OCT-1999:	99US-0158369.
PR	13-OCT-1999:	99US-0159293.
PR	13-OCT-1999:	99US-0159294.
PR	13-OCT-1999:	99US-0159295.
PR	14-OCT-1999:	99US-0159329.
PR	14-OCT-1999:	99US-0159330.
PR	14-OCT-1999:	99US-0159331.
PR	14-OCT-1999:	99US-0159637.
PR	14-OCT-1999:	99US-0159638.
PR	18-OCT-1999:	99US-0159584.
PR	21-OCT-1999:	99US-0160741.
PR	21-OCT-1999:	99US-0160767.
PR	21-OCT-1999:	99US-0160768.
PR	21-OCT-1999:	99US-0160770.
PR	21-OCT-1999:	99US-0160814.
PR	21-OCT-1999:	99US-0160815.
PR	22-OCT-1999:	99US-0160980.
PR	22-OCT-1999:	99US-0160981.
PR	22-OCT-1999:	99US-0160989.
PR	25-OCT-1999:	99US-0161404.
PR	25-OCT-1999:	99US-0161405.
PR	25-OCT-1999:	99US-0161406.
PR	26-OCT-1999:	99US-0161359.
PR	26-OCT-1999:	99US-0161360.
PR	26-OCT-1999:	99US-0161361.
PR	28-OCT-1999:	99US-0161920.
PR	28-OCT-1999:	99US-0161992.
PR	28-OCT-1999:	99US-0161993.
PR	29-OCT-1999:	99US-0162142.

```

Query Match: 54.1%; Score: 867.5; DB: 21; Length: 316;
Best Local Similarity: 56.8%; Pred. No. 9.4e-78;
Matches: 176; Conservative: 45; Mismatches: 82; Indels: 7; Gaps: 3;

OY 10 LVAVSLSCVAHAQLSPREYASSCPNLQSLIVAAAMQAQVASEQMGASLILRFHDCFVQ 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 LMMIMLASQSEAFQNLRDREYKESCPSELFLVYARVYKFAAREPBMGASLLRLRFHHCFFV 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 70 GCDSSILLDAG----GKTAGPNLVSVEGEVIDIRKENVFAACPGVASCADILATAARD 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 GCDSSILLDTPFPLEKTSGPSNNSVYKGFVYLIDIKIKVYKMKCPGLVSCADILATAARD 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 126 GTNLGGPTWSVPLIGRRDSTTASASIANSP-PPPTASLGLTILSEFGQLSPRDMTALS 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 SVLLGGPGMSVVLIGRRDSTTANFAAANSVGLPPPTITLSLIRFAQGLSTRDVAVALS 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 185 GAHTIGARCTTFRGRITGYGTDINASFALRQOTCFRGGGCGN--LAPIVYCTPVEFDDA 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 GSHITGAQCVTFRNKTIYANSNDISFASKRRNCPAISGGGDKKKNLDVRSYDFRDMG 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 243 YFTULSRKQLRHSDDLEFNGSGQDALYKQYSASASLEFNADFAAMIRMGVNYLITGTAG 302
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 FYKQLSKRKILTSDOYLTFNNGPTDLSLVYASHNLNAPYRDFARAMIKMDISPLTGSNG 306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 303 QIRNRCRVVN 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 QIRNCRRPN 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
AACG49113
AACG49113 standard; Protein; 308 AA.

```

XX AAC49113;  
XX 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 62097.  
XX  
XX Protein identification: signal transduction pathway; metabolic pathway;  
KW hydrolisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130049.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132663.  
PR 11-MAY-1999; 99US-0134286.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139839.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 06-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.

```
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 54.0%; Score 866.5; DB 21; Length 308;  
Best Local Similarity 57.3%; Pred. No. 1,1e-77;  
Matches 176; Conservative 43; Mismatches 81; Indels 7; Gaps 3;

```
QY 13 VSLSCVAHAOLSPFFYASCCNLSIVRAMTQAVASEOMGASLRLFFHDCFYOGCD 72
DB 2 IMMLASOSEAOJLRDFFYKESCSFLVVRVYKRAVAREPRMGASLLRFFHDCFYNGCD 61
QY 73 GSTLLDAG---GKTAGFNLNSVGFVVIDTIRKNVEACPGVASCADILALARDGTF 128
DB 62 GSTLLDTPSFLEKTSFGSNNVSGFVVIDKIKFKVEKMGCGIVSCADILALITARDSVL 121
QY 129 LGGPTWSPVPLGRDSTASTASLANSP-PPPTASLGLTSLFSGRGLSPRDMTALSGAH 187
DB 122 LGGPWSVYKLRDSTANFANANSVIPPITLNLNRFKAQGLSTRDMVALSGAH 181
QY 188 TIGQACTFRGRIGYDDTDINAFALROQTCPRSGDGN--LAPIDVOTPVFPTAYFT 245
DB 182 TIGRACQVTFRRIRYNASIDTSFAISKRRNCAPATSGSGDKKANLDVRSFDFHGFYK 241
```

```
QY 246 NLSRRGLFHSDELENGSQDALVRQYSASASLENADEFVAAIMRGVLTGTACQIR 305
DB 242 QLSRKGLTSDOVLFFNNGPTDSLVIAYSHNLNAYRDFARAMIKMGDISPLTGSNGQIR 301
QY 306 RNCRVVN 312
DB 302 QNCRRPN 308
```

Search completed: April 26, 2003, 12:33:32  
Job time : 78 secs



EST

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 08:09:12 ; Search time 1454 Seconds

(without alignments)  
15081.629 Million cell updates/sec

Title: US-10-047-825-3

Perfect score: 1354  
Sequence: 1 aatcgcgcgcgcgttaagc.....aaaaaaaaaaaaaaaaaa 1354

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:  
1: em\_estda:  
2: em\_esthum:  
3: em\_estin:  
4: em\_estmu:  
5: em\_estov:  
6: em\_estpl:  
7: em\_estro:  
8: em\_hlc:  
9: gb\_est1:  
10: gb\_est2:  
11: gb\_hlc:  
12: gb\_est3:  
13: gb\_est4:  
14: gb\_est5:  
15: em\_estfun:  
16: em\_estom:  
17: gb\_gss:  
18: em\_gss\_hum:  
19: em\_gss\_inv:  
20: em\_gss\_pln:  
21: em\_gss\_vrt:  
22: em\_gss\_fun:  
23: em\_gss\_mam:  
24: em\_gss\_mus:  
25: em\_gss\_other:  
26: em\_gss\_pro:  
27: em\_gss\_tod:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	605.6	44.7	848	9	AA979912 MEST3-D6.
2	537.6	39.7	850	9	AA979912 MEST3-D6.
3	537	39.7	633	10	AA979912 MEST3-D6.
4	517.2	38.2	625	10	AA979912 MEST3-D6.
5	500.6	37.0	538	12	AA979912 MEST3-D6.
6	500.6	37.0	602	12	AA979912 MEST3-D6.

7	499.6	36.9	593	10	BE362200	BE362200	DG1_85_CO
8	496.6	36.7	591	10	BE362356	BE362356	DG1_86_CO
9	493.2	36.4	615	10	BE361528	BE361528	DG1_81_G1
10	489.6	36.2	591	12	BE362797	BE362797	DG1_45_F1
11	461.4	34.1	589	10	BE597430	BE597430	DG1_45_F1
12	442.4	32.7	561	10	BE361738	BE361738	DG1_45_F1
13	442	32.6	570	10	BE358471	BE358471	DG1_30_A0
14	439.6	32.5	446	12	BE361858	BE361858	DG1_30_A0
15	437.6	32.3	564	10	BE600142	BE600142	DG1_79_G1
16	437	32.3	555	10	BE671012	BE671012	DG1_79_G1
17	436.6	32.2	560	10	BE600110	BE600110	DG1_79_G1
18	436.6	32.2	566	10	BE366224	BE366224	DG1_31_A0
19	435	32.1	566	10	BE600823	BE600823	DG1_31_A0
20	431.6	31.9	533	10	BE125306	BE125306	DG1_18_A1
21	426.6	31.5	569	10	BE357554	BE357554	DG1_21_H1
22	414	30.6	537	10	BE595535	BE595535	DG1_56_H0
23	413.8	30.6	555	12	BE463190	BE463190	DG1_47_E1
24	410.8	30.3	524	10	BE597461	BE597461	DG1_47_E1
25	406.8	30.0	522	10	BE594365	BE594365	DG1_32_E1
26	385.8	28.5	495	10	BE600529	BE600529	DG1_32_E1
27	384.8	28.4	487	10	BE599526	BE599526	DG1_32_E1
28	381.6	28.2	452	12	BE649323	BE649323	DG1_78_A0
29	375.8	27.8	480	10	BE366914	BE366914	DG1_42_B1
30	374.2	27.6	829	12	BE263468	BE263468	DG1_42_B1
31	372.6	27.5	1303	11	AY107804	AY107804	DG1_42_B1
32	371.8	27.5	476	13	BM324451	BM324451	DG1_31_D
33	370.8	27.4	479	10	BE601026	BE601026	DG1_31_D
34	361.4	26.7	744	14	BQ744368	BQ744368	DG1_31_D
35	360.2	26.6	480	10	BE600167	BE600167	DG1_31_D
36	353	26.1	434	10	BE287460	BE287460	DG1_31_D
37	352	26.0	568	10	BE404079	BE404079	DG1_31_D
38	349.6	25.8	645	10	BE351662	BE351662	DG1_31_D
39	349	25.8	677	12	BE265184	BE265184	DG1_31_D
40	337	24.9	930	10	BE412497	BE412497	DG1_31_D
41	335.8	24.8	601	12	BE462873	BE462873	DG1_31_D
42	332.8	24.6	637	10	BE357660	BE357660	DG1_31_D
43	329.8	24.4	598	10	BE599361	BE599361	DG1_31_D
44	327.8	24.2	815	12	BE267871	BE267871	DG1_31_D
45	327.2	24.2	700	9	AL503342	AL503342	DG1_31_D

## ALIGNMENTS

RESULT 1  
LOCUS AA979912 848 bp mRNA lineat EST 26-MAY-1998  
DEFINITION MEST3-D6.TW1412.Seq ISUM2 Zea mays cDNA clone MEST3-D6 5', mRNA  
ACCESSION AA979912  
VERSION AA979912  
KEYWORDS GI:3157290  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 848)  
Wen,T.J., Ashlock,D.A. and Schnable,P.S.  
Expressed Sequence Tags from B73 Maize Seedlings  
Unpublished (1997)  
Contact: Schnable, PS  
Schnable Laboratory  
Iowa State University  
GA05 Agronomy, Ames, IA 50011, USA  
Tel: (515) 294-0975  
Fax: (515) 294-2299  
Email: schnable@iastate.edu  
PCR Primers  
FORWARD: TW1412 (5'-GAAGATACCCACCAACC-3')  
BACKWARD: TW1412 (5'-TAATACGACTCTATAGGC-3')  
Plate: MEST3 row: D column: 6  
Seq primer: TW1412 (5'-GAAGATACCCACCAACC-3').

# FEATURES

source  
Location/Qualifiers  
1. 848  
/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="MEST3-D6"  
/clone\_1lb="ISUM2"  
/tissue\_type="above ground tissues"  
/dev\_stage="Two-leaf stage green seedling"  
/lab\_host="XLI-MFR Blue"  
/note="Organ: green seedlings; Vector: PAD-GAL4; Site: 1; EcORI: site 2; XhoI: ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with an XhoI oligo-dT primer. The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcORI adaptors, the ds-cDNAs were digested with XhoI and size-selected. The resulting molecules were directionally cloned into the EcORI and XhoI sites of the HybriZAP lambda vector (Stratagene) and excised as PAD-GAL4 phagemids."

BASE COUNT 162 a 284 c 236 g 147 t 19 others

ORIGIN

Query Match 44.7%; Score 605.6; DB 9; Length 848;  
Best Local Similarity 95.7%; Pred. No. 1.1e-81;  
Matches 650; Conservative 0; Mismatches 26; Indels 3; Gaps 3;

325 CGGAACCTGAACTCGCTGCGGCTTTGAGTCAATCGACACCATCAAGCGAAGCTGAG 384  
1 CGGAACCTGAACTCGCTGCGGCTTTGAGTCAATCGACACCATCAAGCGAAGCTGAG 60  
385 GCGCGCTGCGGCTGCTGCTGCGGCGGCGGACATCTCGGCTTGGCGGCGGCGAGGA 444  
61 GCGCGCTGCGGCTGCTGCTGCGGCGGCGGACATCTCGGCTTGGCGGCGGCGAGGA 120  
445 ACCAACCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 504  
121 ACCAACCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 180  
505 GCGAGCGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 564  
181 GCGAGCGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 240  
565 ATCTCCCTGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 624  
241 ATCTCCCTGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
625 CACACCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 684  
301 CACACCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
685 ATCAACGCGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 744  
361 ATCAACGCGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420  
745 AACCTGCGCGCCATCGAGCTGCGAGCGCGGCTGAGGTTGACAGCGGCTTACTTACCAAC 804  
421 AACCTGCGCGCCATCGAGCTGCGAGCGCGGCTGAGGTTGACAGCGGCTTACTTACCAAC 480  
805 CTGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 864  
481 CTGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
865 GAGCGCGCTGAGGAGTACAGCGCGAGCGGCTGCTTCAAGCGGCGGCTTCTGCGCA 924  
541 GAGCGCGCTGAGGAGTACAGCGCGAGCGGCTGCTTCAAGCGGCGGCTTCTGCGCA 599  
925 GCCATGATTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 984  
600 ACCATGATTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 657  
985 AACTGCGGCGGCTGTCACA 1003

Db 658 AACTGCGGCGGCTGTCANAAA 676

||||| |

RESULT 2  
A1374530/C 850 bp mRNA linear EST 21-JAN-1999  
LOCUS A1374530  
DEFINITION MEST3-D6. POLYT-N. Seq ISUM2 Zea mays cDNA clone MEST3-D6 5', mRNA  
sequence.  
ACCESSION A1374530  
VERSION A1374530.1 GI:4174550  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 850).  
Wen, T.J., Ashlock, D.A. and Schnable, P.S.  
Expressed Sequence Tags from B73 Maize Seedlings  
Unpublished (1997)  
JOURNAL Contact: Schnable, PS  
COMMENT Schnable laboratory  
Iowa State University  
6405 Agronomy, Ames, IA 50011, USA  
Tel: (515)-294-0975  
Fax: (515)-294-2299  
Email: schnable@iastate.edu  
PCR Primers  
FORWARD: tw1412 (5'-GAAGATACCCGACCAAC-3')  
BACKWARD: T7-10 (5'-TAATACGACTCATATGAGG-3')  
Plate: MEST3 Row: D Column: 6  
Seq primer: POLYT-N (5'-TTTTTTTTTTTTTTTTTTTTTTT(Agc)-3').  
Location/Qualifiers  
1. 850  
/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="MEST3-D6"  
/clone\_1lb="ISUM2"  
/tissue\_type="above ground tissues"  
/dev\_stage="Two-leaf stage green seedling"  
/lab\_host="XLI-MFR Blue"  
/note="Organ: green seedlings; Vector: PAD-GAL4; Site: 1; EcORI: site 2; XhoI: ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with an XhoI oligo-dT primer. The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcORI adaptors, the ds-cDNAs were digested with XhoI and size-selected. The resulting molecules were directionally cloned into the EcORI and XhoI sites of the HybriZAP lambda vector (Stratagene) and excised as PAD-GAL4 phagemids."

BASE COUNT 168 a 226 c 210 g 194 t 52 others

ORIGIN

Query Match 39.7%; Score 537.6; DB 9; Length 850;  
Best Local Similarity 86.6%; Pred. No. 1.6e-71;  
Matches 652; Conservative 0; Mismatches 92; Indels 9; Gaps 7;

529 AACCCCG 588  
754 AACCTCCG 695  
589 CTGTCG 645  
694 CTGTCG 635  
646 TGCAACACCTTCCG 703  
634 GTGAACCCCTTCCG 575







QY 797 TCACCACTCTCTGTCGCGGCGGCTGTTCACACTCGGACAGAGACTCTTCAACGCGC 856  
Db 121 TCACCACTCTCTGTCGCGGCGGCTGTTCACACTCGGACAGAGACTCTTCAACGCGC 180  
QY 857 GGTGCGACGAGCGCTGTGAGGAGTACAGAGCGGCGCTGCTTCAACGCGC 916  
Db 181 GGTGCGACGAGCGCTGTGAGGAGTACAGAGCGGCGCTGCTTCAACGCGC 240  
QY 917 TCGTGGACGACATGATTAGATGGGCAACGTTGGGTGTCTACCGGCGCGGAGAGA 976  
Db 241 TCGTGGACGACATGATTAGATGGGCAACGTTGGGTGTCTACCGGCGCGGAGAGA 300  
QY 977 TCAGCGCACTGCGGCGGTCTCAACAGTATGATGACGACGATGATGATGATGAT 1036  
Db 301 TCAGCGCACTGCGGCGGTCTCAACAGTATGATGACGACGATGATGATGATGATGAT 359  
QY 1037 CTCTAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095  
Db 360 CTCTAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419  
QY 1096 AGACCTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1155  
Db 420 AGACCTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479  
QY 1156 CTATAGATTGAACAGCTGCTATACAGTATGATGATGATGATGATGATGATGATGAT 1214  
Db 480 CTATAGATTGAACAGCTGCTATACAGTATGATGATGATGATGATGATGATGATGAT 538

RESULT 6  
LOCUS BG463106 602 bp mRNA linear EST 20-MAR-2001  
DEFINITION EML\_47\_E12\_b1\_A002 Embryo 1 (EML1) Sorghum bicolor CDNA, mRNA  
ACCESSION BG463106  
VERSION BG463106.1 GI:13390208  
KEYWORDS EST.  
SOURCE sorghum.  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
1 (bases 1 to 602)  
clade: Andropogoneae; Sorghum.  
Reid S.P., Cordomier-Pratt, M.-M., Gingle, A. and Pratt, L.H.  
An EST database from Sorghum: developing embryos  
Unpublished (2000)  
Contact: Cordomier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: JEN REV  
High quality sequence stop: 548  
POLYA-No.

FEATURES  
source  
1..602  
Location/Qualifiers  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Embryo 1 (EML1)"  
/note="Organ: Embryos germinated for 24 hr; Vector:  
pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2:  
EcoRI; The library was made from poly-A RNA in the cloning  
vector Lambda Zap II. Clones to be sequenced were  
prepared by mass excision."  
BASE COUNT 101 a 219 c 192 g 90 t  
Query Match 37.0%; Score 500.6; DB 12; Length 602;

Best Local Similarity 91.3%; Pred. No. 6,8e-66;  
Matches 543; Conservative 0; Mismatches 49; Indels 3; Gaps 1;

QY 240 GCTCAGGCTCTCTTCTTCCACAGCTGCTTCTTCAAGGCTGCGAGATGATCTTCTGCA 299  
Db 8 GCTCAGGCTCTTCTTCTTCCACAGCTGCTTCTTCAAGGCTGCGAGATGATCTTCTGCA 67  
QY 300 CGCGGAGGAGGAGAACACCGCGGCGGAGAACTGAACTGGTGGCGGCTTTGAGGTAT 359  
Db 68 CGCGGAGGAGGAGAACACCGCGGCGGAGAACTGAACTGGTGGCGGCTTTGAGGTAT 127  
QY 360 CGACACATCAAGCGGAGGAGTGAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 419  
Db 128 CGACACATCAAGCGGAGGAGTGAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 187  
QY 420 CCTCGGCTTCCGCGCGCGGAGCAACCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 479  
Db 188 CTTTGGCTTCCGCGCGGCGGAGCAACCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 247  
QY 480 GCTCGGCGGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 539  
Db 248 GCTCGGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 307  
QY 540 CCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 599  
Db 308 GTCGAGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 367  
QY 600 CGACATGACGCGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 659  
Db 368 CGACATGACGCGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 427  
QY 660 CGGCGCATCTTCTGCGGCGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 716  
Db 428 CAGCGCGATCTTCTGCGGCGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 487  
QY 717 GACGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 776  
Db 488 GACGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 547  
QY 777 GAGGTTCGACAGCGGCTTCTTCAACAGCTGCTTCTGCGGCGGCGGCGGCGGCGGCGGCGG 831  
Db 548 GAGGTTCGACAGCGGCTTCTTCAACAGCTGCTTCTGCGGCGGCGGCGGCGGCGGCGGCGG 602

RESULT 7  
LOCUS BE362200 593 bp mRNA linear EST 20-JUL-2000  
DEFINITION Dg1\_85\_c06\_b1\_A002 Dark Grown 1 (Dg1) Sorghum bicolor CDNA, mRNA  
ACCESSION BE362200  
VERSION BE362200.1 GI:9303757  
KEYWORDS EST.  
SOURCE sorghum.  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
1 (bases 1 to 593)  
clade: Andropogoneae; Sorghum.  
Cordomier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt  
L.H.  
An EST database from Sorghum: dark-grown seedlings  
Unpublished (2000)  
Contact: Cordomier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: JEN REV

FEATURES  
TITLE  
JOURNAL  
COMMENT

BASE COUNT 101 a 219 c 192 g 90 t  
Query Match 37.0%; Score 500.6; DB 12; Length 602;

High quality sequence stop: 523  
POLYA-NO.

BASE COUNT	94 a	210 c	184 g	105 t
ORIGIN				

Query Match	36.9%;	Score 499.6;	DB 10;	Length 593;
Best Local Similarity	90.8%;	Pred. No. 9.6e-66;		
Matches 532; Conservative	0;	Mismatches 54;	Indels 0;	Gaps 0

AUTHORS	Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudanan,M. and Pratt,L.H.
TITLE	An EST database from Sorghum: dark-grown seedlings
JOURNAL	Unpublished (2000)
COMMENT	Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmp@atluga.edu
FEATURES	Sequences have been trimmed to exclude PolyA, vector and regions below pired quality 16. The threshold for highest quality sequence is 20.
source	Seq primer: JEN REV High quality sequence stop: 531 POLYA-No.
Location/Qualifiers	1..591 /organism="Sorghum bicolor" /db_xref="taxon:4558" /clone_lib="Dark Grown 1 (DG1)" /note="Organ: 5-day-old dark-grown seedlings; Vector: lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."
BASE COUNT	94 a 210 c 101 g 105 t 1 others
ORIGIN	

Query Match	36.7%;	Score 496.6;	DB 10;	Length 591;
Best Local Similarity	90.6%;	Pred. No. 2,7e-65;		
Matches 529;	Conservative 0;	Mismatches 55;	Indels 0;	Gaps 0;

[illegible]

RESULT	8				
BE362356					
LOCUS	BE362356	591 bp	mRNA	linear	EST 20-JUL-2000
DEFINITION	BE362356	DG1_B6_C05.b1_A002	Dark Grown 1 (DG1)	<i>Sorghum bicolor</i> cDNA, mRNA	
ACCESSION	BE362356				
VERSION	BE362356				
KEYWORDS	BE362356.1	GI:9303913			
SOURCE	EST				
ORGANISM	<i>Sorghum bicolor</i>				

000 CGACATGACGGCGCTGTGGGCGCGCACACCATCGGGCAGGCC 591  
 548 CGACATGACGGCGCTGTGGGCGCGCACACCATCGGGCAGGCC 591

[illegible]

Db 68 CGCCGAGGAGGAGAGACGCGCCGCCCAACCAATTCGCGCGCTTTGAGGTAT 127

QY 360 CGACACATCATAGAGGAGAGCTGTAGAGCCGCTGCCCCGCGCTGTGTGCGGACAT 419

Db 128 CGACACATCATAGAGGAGAGCTGTAGAGCCGCTGCCCCGCGCTGTGTGCGGACAT 187

QY 420 CCTCGCGCTTGGCGCGCGGAGCAACACCTTCTCGCGCGCGGAGCACTGTGAGCTGC 479

Db 188 CTTGCTGCTTGGCGCGCGGAGCAACACCTTCTCGCGCGCGGAGCACTGTGAGCTGC 247

QY 480 GCTCGCGCGCGGAGCACTGTGAGAGCGCGGAGCACTGTGAGAGCACTGTGAGAG 539

Db 248 GCTCGCGCGCGGAGCACTGTGAGAGCGCGGAGCACTGTGAGAGCACTGTGAGAG 307

QY 540 CCGGAGCGCGGAGCACTGTGAGAGCGCGGAGCACTGTGAGAGCACTGTGAGAG 599

Db 308 GTCGAGCGCGGAGCACTGTGAGAGCGCGGAGCACTGTGAGAGCACTGTGAGAG 367

QY 600 CGACATGAGCGCGCTGTGCGCGCGGAGCAACATCGGCGGAGCGCGGAGCACTGTGCG 659

Db 368 CGACATGAGCGCGCTGTGCGCGCGGAGCAACATCGGCGGAGCGCGGAGCACTGTGCG 427

QY 660 CGCGCGCATCTACGCGGAGAGCACTGTGAGAGCGCGGAGCACTGTGAGAGCACT 716

Db 428 CAGCGCATCTACGCGGAGAGCACTGTGAGAGCGCGGAGCACTGTGAGAGCACT 487

QY 717 GAGCTCGCGCGGAGCGGAGCGGAGCACTGTGAGAGCGCGGAGCACTGTGAGAGCGGT 776

Db 488 GAGCTCGCGCGGAGCGGAGCGGAGCACTGTGAGAGCGCGGAGCACTGTGAGAGCGGT 547

QY 777 GAGCTCGAGAGCGGCTGTGAGAGCGGAGCGGAGCACTGTGAGAGCGGAGCGG 820

Db 548 GAGCTCGAGAGCGGCTGTGAGAGCGGAGCGGAGCACTGTGAGAGCGGAGCGG 591

## RESULT 11

BE597430

LOCUS P11.69.D10.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA, EST 18-AUG-2000

DEFINITION mRNA sequence.

ACCESSION BE597430

VERSION BE597430.1 GI:9852503

KEYWORDS EST.

SOURCE sorghum.

ORGANISM Sorghum bicolor

REFERENCE 1 (bases 1 to 589) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Sorghum.

AUTHORS Cordomier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt,L.H.

TITLE An EST database from Sorghum: pathogen-induced plants

JOURNAL Unpublished (2000)

COMMENT Contact: Cordomier-Pratt MM

Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV  
High quality sequence stop: 524

## FEATURES

SOURCE

Location/Qualifiers

1..589

/organism="Sorghum bicolor"

/db\_xref="taxon:4558"

/clone\_lib="Pathogen induced 1 (P11)"

/note="Organ: Anthracnose-infected leaves from two-week-old sorghum plants 48 hr after inoculation;

BASE COUNT 97 a 205 c 174 g 112 t 1 others

ORIGIN

Query Match 34.1%; Score 461.4; DB 10; Length 589;  
Best Local Similarity 87.5%; Pred. No. 5e-60;  
Matches 504; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Vector: pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI; Two-week-old sorghum plants (BTX 623 cultivar) were infected with pathogen (isolate PM421 of *Colletotrichum graminicola*, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptibility to anthracnose disease. The library was made from poly-A RNA in the cloning vector Lambda Zap II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

QY 29 TCATTCACCGAGCGTGCAGGACAGGACGAGCTTGCATGCGCTCCACCTTGATGC 88

Db 14 TCTTAAGCAAGTAGCTTACCGAGTGTGCATGAGGCGGCTCTACTGATGC 73

QY 89 AATGCTGTGCGCGCTTCCCTCTCTGCTGTGCGCGGAGCAAGCTGTGCGCGCAAG 148

Db 74 AATCTGTGTGCGCGCTTCCCTCTCTGCTGTGCGCGGAGCAAGCTGTGCGCGCAAG 133

QY 149 TCTATGCGCTCTCTGCGCGGAGCAAGCTGTGCGCGGAGCAAGCTGTGCGCGCAAG 208

Db 134 TCTATGCGCTCTCTGCGCGGAGCAAGCTGTGCGCGGAGCAAGCTGTGCGCGCAAG 193

QY 209 TCGCAAGTGAAGAGAGATGGCGCTCTGCTGAGGCTCTTCTCCAGACTGCTTG 268

Db 194 TCAGCAATGACAGAGAGATGGCGCTCTGCTGAGGCTCTTCTCCAGACTGCTTG 253

QY 269 TTCAGGCTGAGAGAGATGGCGCTCTGCTGAGGCTCTTCTCCAGACTGCTTG 328

Db 254 TTCAGGCTGAGAGAGATGGCGCTCTGCTGAGGCTCTTCTCCAGACTGCTTG 313

QY 329 ACTGAAGTGGTGGCGGCTTGAAGTCAATGACACCATCAAGCGGAGCGGCG 388

Db 314 ACAGCAATGAGAGAGATGGCGCTTGAAGTCAATGACACCATCAAGCGGAGCGGCG 373

QY 389 CGTCCCGCGCGCTGTGCTGCGCGGAGCAATCTCTGCGGCTTGGCGGAGCGGAGCA 448

Db 374 CGTCCCGCGCGCTGTGCTGCGCGGAGCAATCTCTGCGGCTTGGCGGAGCGGAGCA 433

QY 449 ACTTCTGCGCGGCGGAGCTGTGCGGCTGTGCGGCGGAGCAATCTCTGCGGCTTGGCGGAGCGGAGCA 508

Db 434 ATGCTGCGCGGCGGAGCTGTGCGGCTGTGCGGCGGAGCAATCTCTGCGGCTTGGCGGAGCGGAGCA 493

QY 509 GCGGCTGCTGCGGAGCAAGCAAGCGGCGGCGGAGCAAGCGGAGCAAGCGGAGCAAGCGGAGCA 568

Db 494 GCGGCTGCTGCGGAGCAAGCAAGCGGCGGCGGAGCAAGCGGAGCAAGCGGAGCAAGCGGAGCA 553

QY 569 CCGTGTGCGGAGGAGCGGCTGTGCGGCGGAGCA 604

Db 554 CGCTGTGAGAGGAGCGGCTGTGCGGCGGAGCA 589

## RESULT 12

BE361738

LOCUS DGI\_82.H08.b1\_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA

DEFINITION DGI\_82.H08.b1\_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA

ACCESSION BE361738

VERSION BE361738.1 GI:9303295

KEYWORDS EST.

SOURCE sorghum.

ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC 1 (bases 1 to 561)  
Cordonier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt, L.H.

TITLE An EST database from Sorghum: dark-grown seedlings  
JOURNAL Unpublished (2000)  
CONTACT: Cordonier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@prattuga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV  
High quality sequence stop: 514  
POLYA-No.

FEATURES  
source Location/Qualifiers

1..561  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Dark Grown 1 (DG1)"  
/note="Organ: 5-day-old dark-grown seedlings; Vector: Lambda Zap; Site.1: XhoI; Site.2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."  
BASE COUNT 94 a 195 c 166 g 106 t  
ORIGIN

Query Match 32.7%; Score 442.4; DB 10; Length 561;  
Best Local Similarity 89.5%; Pred. No. 3.6e-57;  
Matches 476; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

OY 52 AGGACACCTTGGCTGCTCCAGCTGATGCAATCCCGTCCGCTTCCTC 111  
DB 30 AGGTGAGAGTGGCGGCTCCACCTGATGCAATCCCGTCCGCTTCCTC 89  
OY 112 CTCTCTGTGTCGCCACGACAGCTTCGCCCAGTCTTATGCGTCTCTCCCAAC 171  
DB 90 CTCTCTGTGTCGCCACGACAGCTTCGCCCAGTCTTATGCGTCTCTCCCAAC 149  
OY 172 CTGAGAGCATGTTGCGGCGCGATGACCGAGCGCTGCGACAGTGAAGAGATGGGC 231  
DB 150 CTGAGAGCATGTTGCGGCGCGATGACCGAGCGCTGCGACAGTGAAGAGATGGGC 209  
OY 232 GCGCTCTGCTAGAGCTCTTCTTCAAGACTCTTCTTCAAGAGCTGCGAGCATGATC 291  
DB 210 GCGCTCTGCTAGAGCTCTTCTTCAAGACTCTTCTTCAAGAGCTGCGAGCATGATC 269  
OY 292 CTCTGAGAGCGGAGGAGAGACGCGGCGGAGACCTGAGCTGCGGCGGCTT 351  
DB 270 CTCTGAGAGCGGAGGAGAGACGCGGCGGAGACCTGAGCTGCGGCGGCTT 329  
OY 352 GAGGTATGAGACATCAAGCAAGCAAGTCAAGCGGCGGCTGCGGCGGCTT 411  
DB 330 GAGGTATGAGACATCAAGCAAGCAAGTCAAGCGGCGGCTGCGGCGGCTT 389  
OY 412 GCGGACATCTGCGGCTTCCGCGGCGGAGCAACCTTCTGCGGCGGCTT 471  
DB 390 GCGGACATCTGCGGCTTCCGCGGCGGAGCAACCTTCTGCGGCGGCTT 449  
OY 472 AGCGTCCGCTGCGGCGGAGCAAGTCAAGCAAGCGGCGGCTTCTGCGGCAAG 531  
DB 450 AAGCTCCGCTGCGGCGGAGCAAGTCAAGCAAGCGGCGGCTTCTGCGGCAAG 509  
OY 532 CCCCAGCCCGGAGCGGCGGAGCAAGTCAAGCAAGCGGCGGCTTCTGCGGCAAG 583  
DB 510 CCCCAGCGGAGCGGCGGAGCAAGTCAAGCAAGCGGCGGCTTCTGCGGCAAG 561

RESULT 13  
LOCUS BE358471 570 bp mRNA linear EST 20-JUL-2000  
DEFINITION DG1.30.A01.D1.A002 Dark Grown 1 (DG1) sorghum bicolor cDNA, mRNA sequence.

ACCESSION BE358471  
VERSION BE358471.1 GI:9300028  
KEYWORDS EST

SOURCE

ORGANISM Sorghum.  
Sorghum bicolor.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC 1 (bases 1 to 570)  
Cordonier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt, L.H.

TITLE An EST database from Sorghum: dark-grown seedlings  
JOURNAL Unpublished (2000)  
CONTACT: Cordonier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@prattuga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV  
High quality sequence stop: 566  
POLYA-No.

FEATURES  
source Location/Qualifiers

1..570  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Dark Grown 1 (DG1)"  
/note="Organ: 5-day-old dark-grown seedlings; Vector: Lambda Zap; Site.1: XhoI; Site.2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."  
BASE COUNT 95 a 198 c 167 g 110 t  
ORIGIN

Query Match 32.6%; Score 442; DB 10; Length 570;  
Best Local Similarity 87.4%; Pred. No. 4.1e-57;  
Matches 484; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

OY 29 TCATTACCGAGCGTGCAGACAGCAGCAAGCTTGGCATGAGCTCCACCTTATGC 88  
DB 17 TCCTTACGAGTAGTGTCAACGAGTGTGAGGTTGCCATGGGCTCTCTACTTATGC 76  
OY 89 AATGCTGTGCGGCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 148  
DB 77 AATGCTGTGCGGCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136  
OY 149 TCTATGCT 208  
DB 137 TCTATGCT 196  
OY 209 TCGCAAGTGAAGAGAGTGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 268  
DB 197 TCGCAAGTGAAGAGAGTGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 256  
OY 269 TTCAAGCTGCGAGGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 328  
DB 257 TTCAAGCTGCGAGGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 316  
OY 329 ACCTGACGTGCTGCGGCTTGTGAGTCAACATCAAGCAAGCGGAGCGGCGG 388  
DB 317 AGCCCAATTCGCGGCGGCTTGTGAGTCAACATCAAGCAAGCGGAGCGGCGG 376  
OY 389 CGTGGCCGCGGCTGTGTCTGCGGCGGACATCTCTCTCTCTCTCTCTCTCTCTCT 448

Db 377 CGTCCCCGGCGTGTCTGCGCGGACATCTGCGCTGCGCGGAGACGAGCA 436  
 Oy 449 ACCCTTGGGGGGGCGACCTGAGAGCGCTGCGCGGAGAGCTGAGAGCGCA 508  
 Db 437 ATCTGCTGGGGGGGCGACCTGAGAGCGCTGCGCGGAGAGCTGAGAGCGCA 496  
 Oy 509 GCGCGCTGCGCGGAGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 568  
 Db 497 GCGCGCTGCGCGGAGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCT 556  
 Oy 569 CCTGTGCGCGAG 582  
 Db 557 CGCTGTGCGAGG 570

RESULT 14  
 BG841858/c 446 bp mRNA linear EST 29-MAY-2001  
 LOCUS MEST27-D09.T3 ISUM4-TN Zea mays cDNA clone MEST27-D09 3', mRNA  
 DEFINITION sequence.  
 ACCESSION BG841858  
 VERSION BG841858.2 GI:14244122  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 446)  
 Olm.F., Gui.F., Guo.L., Ashlock,D.A., Men.T.J. and Schnable,P.S.  
 Expressed sequence tags from B73 Maize Seedlings and Silks  
 Unpublished (2001)  
 On May 25, 2001 this sequence version replaced gi:14208180.  
 CONTACT: Patrick S. Schnable  
 Schnable Laboratory  
 Iowa State University  
 409 Agronomy, Iowa State University, Ames, IA 50011-1010, USA  
 Tel: 515-294-0975  
 Fax: 515-294-2299  
 Email: schnable@iastate.edu  
 PCR PRIMERS  
 FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)  
 BACKWARD: T3 (ATT AAC CCT CAC TAA AG)  
 Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).  
 location/Qualifiers

FEATURES  
 source  
 1..446  
 /organism="Zea mays"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone\_lib="MEST27-D09"  
 /clone\_lib="ISUM4-TN"  
 /tissue\_type="Seedling and silk"  
 /lab\_host="PH10B"  
 /note="Vector: pT73PAC; Site 1: EcoRI; Site 2: NotI;  
 ds-cDNA molecules were generated as follows: First-strand  
 cDNA was prepared from oligo-dT selected mRNA by priming  
 with a NotI oligo-dT primer (5'-  
 AACTGAGACATTCGCGCGGAGAGCTTTTCTTTTCTTTT). The  
 resulting DNA:RNA hybrid was treated with RNase H and used  
 as a template for DNA Polt-catalyzed second strand  
 synthesis. After the addition of EcoRI adaptors, the  
 ds-cDNAs were digested with NotI and site-selected. The  
 resulting molecules were directionally cloned into the  
 EcoRI and NotI sites of the pT73PAC vector. The library  
 then went through one round of normalization to Cot value  
 of 5 based on the methods of Marcelo Bento Soares (Genome  
 Research 6: 791-806, 1996)."  
 Research 6: 791-806, 1996)."  
 BASE COUNT 65 a 147 c 166 g 68 t  
 ORIGIN  
 Query Match 32.5% Score 439.6; DB 12; Length 446;  
 Best Local Similarity 99.1%; Pred. No. 1.le-56;

Matches 442; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 Oy 200 CCCAGGCGCTGCAAGTAGAGAGATGAGCGCGCTCTGCTAGAGCTCTTTCAGC 259  
 Db 446 CCCAGGCGCTGCAAGTAGAGAGATGAGCGCGCTCTGCTAGAGCTCTTTCAGC 387  
 Oy 260 ACTGCTCTGTTCAAGGCTGCGAGAGATGATCTTCTGACCGCGAGAGAGAGC 319  
 Db 386 ACTGCTCTGTTCAAGGCTGCGAGAGATGATCTTCTGACCGCGAGAGAGAGC 327  
 Oy 320 CCGGGCGGAACCTGAACTGGTGGCGCGCTTTGAGCTATGACACCATCAAGCGAAG 379  
 Db 326 CCGGGCGGAACCTGAACTGGTGGCGCGCTTTGAGCTATGACACCATCAAGCGAAG 267  
 Oy 380 TCGAGGCGCGTGCCTGCGCGCTGCTGTCGCGGACATCTGCGCTGCGCGCGC 439  
 Db 266 TAGAGGCGCGTGCCTGCGCGCTGCTGTCGCGGACATCTGCGCTGCGCGCGC 207  
 Oy 440 ACGGACCAACCTTTCGCGGCGCGAGCTGAGAGCTGCGCGCGCGAGCTGCA 499  
 Db 206 ACGGACCAACCTTTCGCGGCGCGAGCTGAGAGCTGCGCGCGCGAGCTGCA 147  
 Oy 500 CGAGGCGCAAGCGCTGCTGCGCAACAGCAACCCCGCGCGCGAGCGAGCTGCGCA 559  
 Db 146 CGAGGCGCAAGCGCTGCTGCGCAACAGCAACCTCCCGCGCGAGCGAGCTGCGCA 87  
 Oy 560 CGCTCATCTCTCTGTCGCGAGCGCGCTGTCGCGCGCGAGCATGACGCGCTGCGG 619  
 Db 86 CGCTCATCTCTCTGTCGCGAGCGCGCTGTCGCGCGCGAGCATGACGCGCTGCGG 27  
 Oy 620 GCGGCGCAACCATGCGCGCGCGCGG 645  
 Db 26 GCGGCGCAACCATGCGCGCGCGCGG 1

RESULT 15  
 BE600142 564 bp mRNA linear EST 18-AUG-2000  
 LOCUS P11\_79.G11.b1.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,  
 DEFINITION mRNA sequence.  
 ACCESSION BE600142  
 VERSION BE600142.1 GI:985215  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Sorghum.  
 1 (bases 1 to 564)  
 Cordumier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt  
 L.H.  
 An EST database from Sorghum: pathogen-induced plants  
 Unpublished (2000)  
 CONTACT: Cordumier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 543 0210  
 Email: mmpratt@uga.edu  
 Sequences have been trimmed to exclude polyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: JEN REV  
 High quality sequence stop: 562  
 POLYA-No.  
 location/Qualifiers  
 1..564  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Pathogen induced 1 (P11)"  
 /note="Organ: Anthracnose-infected leaves from  
 two-week-old sorghum plants 48 hr after inoculation;







GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 09:09:12 ; Search time 130 Seconds  
(without alignments)  
12307.038 Million cell updates/sec

Title: US-10-047-825-3  
Perfect score: 1354  
Sequence: 1 aattggccgcagcgttaagc.....aaaaaaaaaaaaaaaaaaaaa 1354

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues  
Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_MA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	174.2	12.9	1017	9	US-09-938-842A-60 Sequence 60, Appl
2	172.4	12.7	1281	12	US-10-027-559-11 Sequence 11, Appl
3	171.4	12.7	1011	9	US-09-938-842A-2545 Sequence 2545, Ap
4	157.4	11.6	824	10	US-09-894-633A-83 Sequence 83, Appl
5	150	11.1	996	9	US-09-938-842A-1711 Sequence 1711, Ap
6	139.6	10.3	1041	9	US-09-938-842A-1632 Sequence 1632, Ap
7	128.2	9.5	251	10	US-09-923-876-1167 Sequence 1167, Ap
8	127	9.4	1011	9	US-09-938-842A-2559 Sequence 2559, Ap
9	118.2	8.7	234	10	US-09-923-876-3446 Sequence 3446, Ap
10	117.4	8.7	993	9	US-09-938-842A-2693 Sequence 2693, Ap
11	116.2	8.6	1077	9	US-09-938-842A-1980 Sequence 1980, Ap
12	114.8	8.5	1008	9	US-09-938-842A-1714 Sequence 1714, Ap
13	114	8.4	942	9	US-09-938-842A-2367 Sequence 2367, Ap
14	113.4	8.4	981	9	US-09-938-842A-783 Sequence 783, Appl
15	110	8.1	987	9	US-09-938-842A-1518 Sequence 1518, Ap
16	109.8	8.1	1011	9	US-09-938-842A-1719 Sequence 1719, Ap
17	107.4	7.9	966	9	US-09-938-842A-1273 Sequence 1273, Ap
18	105.4	7.6	1059	9	US-09-938-842A-2325 Sequence 2325, Ap
19	103.4	7.6	981	9	US-09-938-842A-1340 Sequence 1340, Ap

c	20	101.8	7.5	1045	10	US-09-770-445-166	Sequence 166, App
	21	100.2	7.4	938	10	US-09-923-876-97	Sequence 97, Appl
	22	99	7.3	275	9	US-09-938-842A-2329	Sequence 2329, Ap
c	23	96	7.1	1014	10	US-09-770-445-217	Sequence 217, App
	24	95	7.0	969	9	US-09-938-842A-2446	Sequence 2446, App
	25	93.2	6.9	373	10	US-09-878-574-710	Sequence 710, App
	26	92	6.8	441	10	US-09-770-444-907	Sequence 907, App
	27	90.6	6.7	318	10	US-09-923-876-3642	Sequence 3642, App
	28	88.8	6.6	285	10	US-09-878-574-8092	Sequence 8092, Ap
	29	87	6.4	239	10	US-09-923-876-2727	Sequence 2727, Ap
	30	82	6.1	966	9	US-09-938-842A-1498	Sequence 1498, Ap
	31	81.6	6.0	1041	9	US-09-938-842A-858	Sequence 858, App
	32	81.4	6.0	1062	9	US-09-938-842A-1457	Sequence 1457, App
	33	81	6.0	755	10	US-09-770-444-989	Sequence 989, App
	34	80.4	5.9	1062	9	US-10-101-736-3	Sequence 3, Appl1
	35	80.4	5.9	1062	9	US-09-834-659-3	Sequence 3, Appl1
	36	80.4	5.9	1062	10	US-09-834-656-3	Sequence 62, Appl1
	37	78	5.8	972	9	US-09-938-842A-682	Sequence 3718, Ap
	38	77.8	5.7	303	10	US-09-923-876-3718	Sequence 1596, Ap
	39	77.8	5.7	960	9	US-09-938-842A-1596	Sequence 2184, Ap
	40	77	5.7	990	9	US-09-938-842A-2184	Sequence 50, Appl
c	41	77	5.7	1181	10	US-09-770-445-50	Sequence 2656, Ap
	42	75.4	5.6	1074	9	US-09-938-842A-2656	Sequence 2630, Ap
	43	74.8	5.5	262	10	US-09-923-876-2630	Sequence 2570, Ap
	44	74.2	5.5	954	9	US-09-938-842A-2570	Sequence 2181, Ap
	45	72.8	5.4	1116	9	US-09-938-842A-2181	

## ALIGNMENTS

RESULT 1  
US-09-938-842A-60  
Sequence 60, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Krepes, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SCRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938, 842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227, 866  
PRIOR FILING DATE: 2000-08-24  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/264, 647  
PRIOR FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: US 60/300, 111  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 60  
LENGTH: 1017  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-60

Query Match 12.9% Score 174.2; DB 9; Length 1017;  
Best Local Similarity 52.28; Pred. No. 6.3e-37;  
Matches 489; Conservative 0; Mismatches 408; Indels 36; Gaps 3;

QY	114	CTCTGTGTGTCGCCACGACGCTCTGCGCCAGCTCTATGCTCTCTGCTGCCCAACT	173
DB	84	CTATGGCGGAGACAAAGGAGAACTCTCCAGAGTTCTACAGAACTCGTCCCTAGAGC	143
QY	174	GCAGAGCATGTCGCGGCGGCGATGACCCAGCGCCGTCGCAAGTGACAGATGGCCG	233
DB	144	CGAGGAGATGATGAGGACAGTGTACCAAGTGTGGCAAGGAGACATGATGGCTGC	203
QY	234	CTCTCTGCTGAGGCTCTCTCTCCAGCAGCTCTGTTCAAGGCTGCGAGATGATGCT	293
DB	204	TTCCTCATGAGCTTCATTCACAGATGTCTTCGTCAGGCTTGATGATGATGCTTGC	263

```

OY 294 TCTGACGCCGGAGG-----GAGAGACGCCGGGCGGAACTGAACCTGCCT 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 TCTAGACACCACTGGAGATATGTTACTGAGAGAACTTAAACCGAAGCAGATCGGC 323
OY 342 GCGCGGCTTTGAGGTCATGACACCATCAAGCGAAGCTGAGGCGGCGGCGGCT 401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 324 TCGCGGGTTGAAGTTGTTGACGAGATCAAGGACAGATTAGAGAAAGATGCCCTTAAC 383
OY 402 CGTGTGTCGGCGGACATCTCGCGCTTGCGCGCGCGGCGGCGGCGGCGGCGG 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 TGTTCCTTGTGTCGATGCTTACCTAGCCCGCTAGAGACTCTCTGTTTACTGCTGG 443
OY 462 GCGGACCTGAGCGTCCGCGCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 444 ACCAAGCTGATGATGTTCTTTGGGAAGAAGATTCGAAAGTCGAACGCTTGAGAGATC 503
OY 522 CAACAGCAACCCCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 504 AAACAGCAACATCTGACACCCAGCAACATCTTCAACATTTGTACAGAGATTTAAACA 563
OY 582 GCAGGCGCTGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 564 CCAAGTCTCGATCTACCGGAGCTGTGCTCTCGGCGGAGCCACACCATTTGATTTTC 623
OY 642 CCGGTGACACACCTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 624 AAGATGCACTAGTTTCAGACAGACTGTACAAACCATCCGGAACGGAATGCCGAGAG 683
OY 681 CGACATCAACGCGCTCTTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 684 AACCTTAGACAACTCTGAGCTGTCTTAACTTGGCCAAAGGTCCTTAATGCGGGGGA 743
OY 741 CGGCAACCTGCGGCCATGAGCTGACAGCGCGCGGTGAGTTCGACAGCGGCTACTTAC 800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 744 CAGAACTGTGAGAGCTTGAATCAAGCGGTGGAAGATTGTAACACTTACTTAA 803
OY 801 CAACCTGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 860
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 804 GAACCTGATCGAAGATGCGGCTGTGTAATTCGACAGGCTGTTCTTACCAATGA 863
OY 861 GCAGGAC---GCGTGTGTGAGGAGCTACAGCGCGGCGGCGGCGGCGGCGGCT 917
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 864 GCAATCAAGAGAGCTGCTGTCAGAAAGTATGACAGAGATCAAAAGATTTCTTGAGCACTT 923
OY 918 CGTGCACCATGATTTAGATGCGCAACGTTGGGTGCTCACCGGACCGCGGACAGAT 977
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 924 CGCGGAGTCGATCAAGATGCGGGAATATCTCTCCCTTGACAGGTTGAGATGGAT 983
OY 978 CAGCGCAACTGCGCGGCTGTCACAGCT 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 984 CAGAAAGATTGCAAGAAATTAATACT 1012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 2
US-10-027-559-11
: Sequence 11, Application US/10027559
: Patent No. US20020144307A1
: GENERAL INFORMATION:
: APPLICANT: Simmons, Carl
: TITLE OF INVENTION: Maize Defense-Inducible Genes and Their
: FILE REFERENCE: 035718/239835
: CURRENT FILING DATE: 2001-10-23
: PRIOR APPLICATION NUMBER: US/10/027, 559
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 1281
: TYPE: DNA
: ORGANISM: Zea mays

```

```

FEATURE:
: NAME/KEY: misc-feature
: LOCATION: (0)...(0)
: OTHER INFORMATION: Peroxidase-like cDNA
: NAME/KEY: CDS
: LOCATION: (18)...(1088)
US-10-027-559-11

```

```

Query Match 12.78; Score 172.4; DB 12; Length 1281;
Best Local Similarity 52.8%; Pred. No. 2,2e-36;
Matches 491; Conservative 0; Mismatches 376; Indels 63; Gaps 3;

```

```

OY 136 CTCTGCCCAAGTCTCTATGAGTCTCTCTGCGGCGGCGGCGGCGGCGGCGGCGG 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84 CTCAGTTCGTTCTTCAACAAAGACGTGCCCCAGGCGGCGGCGGCGGCGGCGGCGG 143
OY 196 ATGACCCAGGCGCTGCAAGTGAAGCAGAGAGAGGCGGCGGCGGCGGCGGCGGCGG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 GTGGCGCGCGGTTCAACAACTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 203
OY 256 CAGGAGTCTGTTCAAGGCTGCGAGCGATGATCTTCTGACGCGGAGGAGGAAG 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 CATGACTGTTCTGAGAGGCTGCGAGGCTGCTGCTGATGATGATGATGATGATGAT 263
OY 316 ACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 AAGCGGAGAGAGAGCTGATCCCAACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 323
OY 364 ACATCAAGCGGAAGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 324 CCGGCGCAAGGCGTCCCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 383
OY 424 GCGGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 GCCTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 443
OY 484 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 444 GAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 503
OY 544 ACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 504 TTTCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 563
OY 604 ATGACGCGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 564 ATGTTGTCTCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 623
OY 664 CGCATCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 624 ATTAAACAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 683
OY 695 -----CTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 684 GCGCTGAGCAAAAGCTTACGCGTCTCTCTCAAGAGATTTGCGGCGGCGGCGGCGGCGG 743
OY 738 -----CGAGGCAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 744 TTTCTCCCAACACGAGGAGTTCATGAGCTCATGACCGCGGCGGCGGCGGCGGCGGCGG 803
OY 793 TACTTACCACTGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 804 TACTAGTGGGCTTACCAACAACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 863
OY 853 GCGGCGTCCAGAGCGGCTGTGAGAGCGTACAGGCGGCGGCGGCGGCGGCGGCGGCGG 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 864 AACGCAACATGAAGCGCTGTGAGAGCTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 923
OY 913 GACTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 924 AAGTTCGCAAGTCATGCTCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 983
OY 973 CAGATAGGCGCAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1002
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

      Db      984 GAGATCAGCGCCACTGCAGGGTCATCTACAC 1013
      RESULT 3
      US-09-938-842A-2545
      : Sequence 2545, Application US/09938842A
      : Patent No. US20020160378A1
      : GENERAL INFORMATION:
      : APPLICANT: Harper, Jeff
      : APPLICANT: Kreps, Joel
      : APPLICANT: Wang, Xun
      : APPLICANT: Zhu, Tong
      : TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
      : FILE REFERENCE: SAME, AND METHODS OF USE
      : FILE REFERENCE: SCRIPI300-3
      : CURRENT APPLICATION NUMBER: US/09/938,842A
      : CURRENT FILING DATE: 2001-08-24
      : PRIOR APPLICATION NUMBER: US 60/227,866
      : PRIOR FILING DATE: 2000-08-24
      : PRIOR APPLICATION NUMBER: US 60/264,647
      : PRIOR FILING DATE: 2001-01-16
      : PRIOR APPLICATION NUMBER: US 60/300,111
      : PRIOR FILING DATE: 2001-06-22
      : NUMBER OF SEQ ID NOS: 5379
      : SEQ ID NO 2545
      : LENGTH: 1011
      : TYPE: DNA
      : ORGANISM: Arabidopsis thaliana
      : US-09-938-842A-2545
  
```

Query Match	12.7%	Score 17.1	DB 9	Length 1011
Best Local Similarity	52.1%	Pred. No. 3.6e-36		
Matches 481, Conservative	0	Mismatches 406	Indels 36	Gaps 3

Db	617	GGTGCACAAAGTTTCGACAGAGAGTTTATATACATCAATCCGGCACAAGGAACCTGTATATGA	676
QY	686	---TCACGCGCTCCTTCGCGCGCTGCGGACAGACAGACGTGCCGGTCCGTCGGCGAGC	742
Db	677	CTCTTAGGCAATTAATATGCAACACATTAATGGCTCAACGATGTCCGAGATCCGAGAGGTACC	736
QY	743	GCAACCTGGGCGCCATCGAGAGCTGGAGACGCCGGTAGGTTGACACAGCGCTTACTTACCA	802
Db	737	AGACCTTAATTTCTCCTCGACTTCGCGACACCGTTCAAAGTTTGAACACACTTACTTACAGA	796
QY	803	ACCTCTCTWCGCGCGCGGCGCTTCTTCACATCGACACGACGAGCTTTTCACGCGCGGTCCG	862
Db	797	ACCTGATTAATGTACAAAGGTATTATTAGAGCTGTGATGAGATTCTGTCTTACGAAAGACAAC	856
QY	863	AGGAC---GCGCTGGTGAGGCGAGTACAGCGCCAGCGCGCTCTTTCACGCCGCACTTG	919
Db	857	AGTCCAAAGAGACTGGTGGAGCTTAACGCTGAGATCAAGAGCGCTCTTTAGACAGTTTG	916
QY	920	TGGCAGCCATGTTTNGAATGGGCAACGTTGGGGTCTACCCGGACACCGCCGAGACATCA	979
Db	917	CTAAATTCATGGTGAAGATGGGAATATATCTCTCCGTTGACAGGCGCCAAAGGAGAGATCA	976
QY	980	GGCGCACTGCCGGGCTGTCAAC	1002
Db	977	GCGGTATCTGTGAGAGGCTTAAAC	999

RESULT 4  
US-09-894-633A-83  
; Sequence 83, Application US/09894633A

```

: GENERAL INFORMATION:
: APPLICANT: Conner, Timothy
: APPLICANT: Dubois, Patrice
: APPLICANT: Malvern, Marianne
: APPLICANT: Masucci, James
: TITLE OF INVENTION: PLANT REGULATORY SEQUENCES FOR SELECTIVE CONTROL OF GENE EXPRESSION
: FILE REFERENCE: 38-21(15856)B
: CURRENT APPLICATION NUMBER: US/09/894.633A
: CURRENT FILING DATE: 2001-06-28
: PRIOR APPLICATION NUMBER: 60/214,357
: PRIOR FILING DATE: 2000-06-28
: PRIOR APPLICATION NUMBER: 09/894,633
: PRIOR FILING DATE: 2000-06-28
: NUMBER OF SEQ ID NOS: 111
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 83
: LENGTH: 824
: TYPE: DNA
: ORGANISM: Zea mays
US-09-894-633A-83

Query Match      11.6%   Score 157.4; DB 10; Length 824;
Best Local Similarity 54.0%   Pred. No. 1.9e-32;
Matches 425; Conservative 0; Mismatches 311; Indels 51; Gaps 3;

```

Db	317	GTGGTTTGGAACTCATGTGAGGAAATCAAACTGGCTTAGAACAAGATGTGCTGAAACAG	376
Qy	404	TGTGTGCGCCGACATCTCTGCGCTTTGCGCGCGCGACGAGAACCACTTTCTGGGGGCG	463
Db	377	TTTTTGGCGTAATCTTTGGCTTAACCGGTAAAGACTCAACTGTATTACGGGTGAC	436
Qy	464	CGAGCTGGAGGTCGCGCTGGGGGGGAGACTGAGACGGGCGGCTGCTCGCCA	523
Db	437	CGAGCTGGGAGTACCTCTAGGAAGAAAGACGCGAAGGAGCAAGCTTAGTGGTTCCA	496
Qy	524	ACAGCAACCCCCCGCCCGACGGCGCAAGCTCGGCAAGCTATCTCCTGTTGGCGAGCG	583
Db	497	ACAAGGAAATTCGCTGCCAAACAACACTTTCCAACCAATCTCTACATAAATGAAGCTC	556
Qy	584	AAGGCGTCTGCGCCGCGACATGACGGCGGTGTGGGCGCGACACCATGGGCGAGGCC	643
Db	557	AAGGCGTTGATCTGTGTGATCTTCTCTCCCTCTAGGAATGCACCATTTGGAAACTCGA	616
Qy	644	GGTGCACCACTTCGCGGCGCGCATCTACGGCGCACATCGACA-----	685

Query Match	Similarity	11.6%	Score 157.4	DB 10	Length 824
Best Local	Similarity	54.0%	Pred. No. 1, 9e-32		
Matches	425	Conservative	0	Mismatches	311
				Indels	51
				Gaps	
Qy	112	CTCTCTGTGTGCCCGCACAGCATGCTCTGCGCCAGCTTCTATGCTGCTCTCTGCGCCAAC	171		
Db	7	CTGGGCGGCGGGGGCGGGCGGCCCTTGAGCGCGGACTTCTATGCCAGTGTGCGCCCGC	66		
Qy	172	CTGCAGAGCATGTTGCGGCGCGGATGATGCCAGGCGCGTCGCAAGTGAAGACAGATGGG	231		
Db	67	GGGAGCGGATCATTTGCGGAGGTATGACAGACGAAGACAGATGGGGAACCGACAGCGGCC	126		
Qy	232	GCCTCTGTGCTCAGGCTCTTCTCCACAGATGCTTTCGTTCAAGCTGCGAGATGCATC	291		
Db	127	GGGGGCTGCTCTCCGGTCTTCTTCCACACATGCTTTCGTCACGCGGGTGCAGCGCTCGTG	186		
Qy	292	CTTTCGAGCGCGGAGGAGAGAGACCGCGCGGCGGAACCTGAA-----C	336		
Db	187	CTGATCGGCTCCACCCAGTTTCCAGAAAGTCGAGACACGCGGAGATCAATCACTCCCTC	246		
Qy	337	TCGGTCGGCGGCTTTGAGAGTCATGCACCATCAAGCGGAACGTGAGCGCGGCTGCCCC	396		

Db	247	CCCCGGGAGACGCTTGGACGCGCTGGTGGCGGCGCAAGCTGGCCCTGGAGCTGGAGTGGCCCC	306
OY	397	GCGCTGTGTGTGCTGCGCGCGACATCTCTGCGCTTGGCCGCGCGACGAGCAACCACTTCTC	456
Db	307	GGGGTGTGTCTCTGCGCGCGACATCTCTGCACTGGCGGTGGGCGTGTGATTACATGAC	366
OY	457	GCGCGCGGACACTGGAGCGCTGGCGTGGGGGGGGAGCTGGAGAGCGCCAGCGGCTCG	516
Db	367	GCGGGGGCCCCGGTACCGCGTTCCGTTGGGGGCGAGAGACTGCTGTCTGTCTGCGCCAG	426
OY	517	CTCGGCAACAGCAACCCCGCCCGCGAGCGGCGACGCTCGGCGACGTCATCTCCCTGTTTC	576
Db	427	GCGCGCGAGTGGAGCTCGCGAGCCCAATTCAACCGTGGAGACCGGCTATCCAGATGTTTC	486
OY	577	GCGAGGAGGCGCTGTCTGCGCGCGGACATGACGGCGCTGTGGGGCGGCCAACCATCGGG	636
Db	487	GCGCGCAAGGGGTTCTACGCTGCAGGAGCTGGTGGGCGCTGCCGGCCACAGCGTGGGC	546
OY	637	CAGGCGCGGTGCACACCTTCCGCGGCGGCAGTCTCAGG-----	674
Db	547	TTCTTCCACTGCGAGAGATTTGGCGGACCGCCTCTTCAACAATTCGCAACACGAGGCGGAG	606
OY	675	-----CGACACGCAATCAAGCTCTTTCGCGGCGCTGCGGCGACAGACGTGC	723
Db	607	CCGAGACAGTTGAGCCCGACAGCATGAACCGGCTCTACGCGACAGGGGCTTCAGGAGCTGTC	666
OY	724	CCGCGGTCCGGCGGCGAGCGCAACTGGCGCC---CATGCAAGTGCAGACGCCGCGTGAGG	780
Db	667	AGGGACTCTCCTCAAGAGACCCACATCGCGCGGTTCAACGACATCATATACGCCCGGGCAG	726
OY	781	TTTCGACAGGCGCTACTTTCACCAACCTGTGTGTGCGGCGGCGCTGTTCACATGGAGCAG	840
Db	727	TTTCGACACAGTACTTCTGTCAACTCGAGGGCGGCGCTCTGAGCAGCAGCAG	786
OY	841	GAGCTCT 847	
Db	787	GAGCTGT 793	

RESULT 5  
 US-09-938-842A-1711  
 Sequence 1711, Application US/09938842A  
 Patent No. US20020160378A1  
 GENERAL INFORMATION:  
 APPLICANT: Harper, Jeff  
 APPLICANT: Kreps, Joel  
 APPLICANT: Wang, Xun  
 APPLICANT: Zhu, Tong  
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 TITLE OF INVENTION: SAME, AND METHODS OF USE  
 FILE REFERENCE: SCRIP1300-3  
 CURRENT APPLICATION NUMBER: US/09/938, 842A  
 CURRENT FILING DATE: 2001-08-24  
 PRIOR APPLICATION NUMBER: US 60/227,866  
 PRIOR FILING DATE: 2000-08-24  
 PRIOR APPLICATION NUMBER: US 60/264,647  
 PRIOR FILING DATE: 2001-01-16  
 PRIOR APPLICATION NUMBER: US 60/300,111  
 PRIOR FILING DATE: 2001-06-22  
 NUMBER OF SEQ ID NOS: 5379  
 SEQ ID NO 1711  
 LENGTH: 996  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-1711

	Query Match	11.1%;	Score 150;	DB 9;	Length 996;	
	Best Local Similarity	50.3%;	Pred. No. 2e-10;			
	Matches 477; Conservative	0;	Mismatches 435;	Indels 36;	Gaps 3;	
OY	91 TGCGTGGTGCCGCTTCCTCCTCTCTGTGTGCCCCACGCACAGCTCTTGCCCCACGTC	150				

Dd	43	TGTTTTGCCCTCTCTGTCTCTGTGGAGCAAGAAGCATATGGAAGGCAAACTTTTCCCGGGTAT	102
OY	151	TATGCGTCCTCCTGCCCCAACCTTGACAGCATCGTTGGGGGGCGGATGACCACGCCTC	210
Dd	103	TACGGCCATTATGATGCCCAAGTCAATGATGATCTGAGATCATAGTTGACTTAAGCTGTT	162
OY	211	GCAAGTGAAGAGAGATGGGGCCCTCTGCTAGGCTCTTCTTCCACGACTGCT TGGTT	270
Dd	163	GCTAAGAGACACGATATGGCTGCTCTCTGTTGTGAGACTTCAATTCCACGACTGTTCGTT	222
OY	271	CAGGCTGCGAGCGATCGATCTTCTCGAGCGCGGAG-----GGAGAGACC	318
Dd	223	CAGGTTGTGATGGCTCTTCTCTTAGACAGCATGGAGGGTAGGACTGTGAGAAAAAC	282
OY	319	GCCGGGCGGCACTGAATCGGTGCGGGCTTTTAGTTCATCGACACCATTAAGCGGAAC	378
Dd	263	TCAACCCCTTAACGAATATCAGCTCTGTGGTTTACGTAGTGCATAATCAAAGCTGAA	342
OY	379	GTCGAGGCCGCGTGCCTCCGCGGTCTGTCTGTGCGCGCACATCTCGCGTTCGCGCGGC	438
Dd	343	CTGGAGAAACAATGCCCTGGAACTGTTCTTGCGCTGATGTTCTTACCCTTAGCGCTAGA	402
OY	439	GACGGAAACACACTTCTTCGGGGGGCCACCGAGAGCGTGGCGCTGGGGGGGGGAGCTCG	498
Dd	403	GACTCTCTGTTCTTACCGGTGACCAAGTTGGTTGTTCANTGGGAABAAGATTC A	462
OY	499	ACGAGCGCCACGCGCTCGCTGCCACAGCAACCCCGCGCCCGCGAGGCGACCTCGGC	558
Dd	463	AGAAATGCAACTTGAGTCAATCGAACAAACATCCCTGCACCAACACACTTTCAG	522
OY	559	ACGCTCATCTCCCTGTCGGAGGACAGGCGCTGCGCGCGGCGACATGAGGGGGCTGCG	618
Dd	523	ACCATCTATCTAATGATTTAAACGCTCAAGGACTTGATATCACTAGACTTGTGCGTCTCTCC	582
OY	619	GCGCGGCAACCATCGGGGACGCGCGGTGACACACTTTCGCGGCGCGCATCTAGCGGAC	678
Dd	583	GGTAGTCAACAAATCGATTCGATGATGACAGAGTTTCAAGACGAGGTTGTACAAACAG	642
OY	679	ACCGACA-----TCAGCGCTCTTCCGCGGGGCTGCGGCGACAG	717
Dd	643	TCCGGAAGGCGAGTCCAGACATGACATTTGAAACAAATCTTCGTGTAACTTGGCGCAA	702
OY	718	ACGTCGCCGCGGTGCGGGGCGAGAGCAACTGTGGCGCCATCGACGTGACAGCGCGTG	777
Dd	703	CGGTGTCAAAATCGGGTGGGACCAAGTTCTCTCGGTGCTAAGACATCATCAGCGCGCG	762
OY	778	AGGTTGACACGCGCTACTTACCACACTGTCTGTGCGGCGGGGCGCTGTTCCACTGGAC	837
Dd	763	AGCTTCGACAACAGCTATTCTAAACATTGATMGAAACAAGAGGGTGTGTAACCTGGAC	822
OY	838	CAGGAGCTTTCACACGG---CGGGTCGAGAGCGCGGTGGTGAAGGCGTACAGCGCAG	894
Dd	823	CAGATTCTGTTACAGCAGTAAACGAGAAAATCGAGAGAGCTTGTGAAGAAGTATGCAAGAT	882
OY	895	GCTCGCTCTTCAACGCCGCACTTGTGTGCGACGCCATGATTTAGATGGGCAACGTTGGGGTG	954
Dd	883	CAAGAGAGATTTTTTGAAGCAATTTGGCGAATTCGATGATCAAGATGGGAAATATCTCC	942
OY	955	CTCACCGGACACCGCGGAGATAGCGGCAACTGCGCGGTGCTCAAC	1002
Dd	943	TTGACAGGTTCTAGTGGGAATAGGAAAGAAATTTGACGGAAAGATTTAAC	990

```

RESULT 6
US-09-938-842A-1632
; Sequence 1632, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

```

```

; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1632
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1632

```

Query Match 10.3%; Score 139.6; DB 9; Length 1041;  
Best Local Similarity 50.7%; Pred. No. 1.3e-27;

Matches 448; Conservative 0; Mismatches 394; Indels 42; Gaps 3;

```

QY 148 TTCTATGCGCTCTGCTGCGCCCAACCTGAGAGATGTTGCGGGCGGATGACCCAGGCC 207
DB 142 TTCTATGAGCGGCTCTGCTGCGCCCTTTCAACGATGTTAAAGTCGGAGTTGGAGAGCT 201
QY 208 GTGCGAAGTGAGCAGAGATGCGGCGCTCTGCTGAGGCTCTTCTTCACGACTGCTTC 267
DB 202 TTTAAAGATGATCTCGAATCGCGCATCTCTTCGACTCCATTTCCACGATGTTGTTT 261
QY 268 GTTCAAGGCTGCGAGGATCGATCCCTCTCGACG-----CGAGGGAGAG 315
DB 262 GTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 321
QY 316 ACCGCGGCGCGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
DB 322 AAGCTCAGCGCAACCGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 381
QY 376 AAGCTGAGGCGCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435
DB 382 GATATGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441
QY 436 CGGAGGAGAACCAACCTTCTGCGGGCGGCGGAGCTGAGGCTGCGGCGGCGGCGGCGG 495
DB 442 AGGAAAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 501
QY 496 TCGACGACGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 555
DB 502 TCATTAAAGCGGAGTGAGCAAGGGGAGATTAACAATCTCCATCTCCGTTTAAAGCGGTT 561
QY 556 GCGACGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 615
DB 562 GAGAAATATACAGCCCAAGTCTGAGACCTTGAGACTGACCTCAAGACGTTGTTCTCTCT 621
QY 616 TCGGCGGCGAGCAACATCGGGGAGGCGCGGTG-----CAACAC 654
DB 622 TCGAGGACACATCTATAGAGATTGCTCAATGTTTCTGATATACAGATAGACTCTTCAAC 681
QY 655 TTCCGCGGCGGATCTACGCGCAACACGATCAACGCTCTCTCTCTCTCTCTCTCTCTCT 714
DB 682 TTCAAGGCTCAAGGCGAGCGTCAAGCCCAACCTAGCGGCTCTCTCTCTCTCTCTCTCT 741
QY 715 CAGACTGCGCGGCTGCGGGCGGCGAGCG-----AACCTGCGGCGCATCAAGCTG 765
DB 742 CTAAAGGACAGCTGCTCAAGCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 801
QY 766 CAGACGCGCGGATGCTTCAACGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 825
DB 802 GCTAGCTCAGTCAAGTTTACATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 861
QY 826 TTCCACTGCGACGAGGCTCTTCAACGCGGCGGCTGCGAGGCGCGCTGCTGAGGAGTAC 885
DB 862 TTGATTTCTGATCAACCTTAATGACAGATCTACGGGCTGCGGCGCTTGTGTAAGTCGTA 921

```

```

QY 886 AGCGCCAGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 945
DB 922 AGCGAGAAATCGCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 981
QY 946 GTTGGGCTCTACCGCGGCGGCGGAGACATCAAGCGCAACTG 989
DB 982 ATCGAGATTATGACCGGAGATGATGAGTAAATGAGAGAAATG 1025

```

RESULT 7

US-09-923-876-1167  
Sequence 1167, Application US/09923876  
Patent No. US20020013958A1  
GENERAL INFORMATION:  
APPLICANT: Lalgudi, Raghunath V.  
APPLICANT: Kamigaki, Laura Y. (Itu)  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING  
FILE REFERENCE: PL-0012-1 CON  
CURRENT APPLICATION NUMBER: US/09/923,876  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: 09/238,329  
PRIOR FILING DATE: 1999-04-21  
PRIOR APPLICATION NUMBER: 60/085,331  
PRIOR FILING DATE: 1998-05-05  
NUMBER OF SEQ ID NOS: 6332  
SOFTWARE: PERL Program  
SEQ ID NO 1167  
LENGTH: 251  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. US20020013958A1 700158231H1  
US-09-923-876-1167

Query Match 9.5%; Score 128.2; DB 10; Length 251;  
Best Local Similarity 72.3%; Pred. No. 7.8e-25;  
Matches 180; Conservative 0; Mismatches 68; Indels 1; Gaps 1;

```

QY 308 GGGAGAAACCGCGCGCGCAACCTGAACTGCTGCGCGGCTTGTGAGTATGACACCA 367
DB 1 GCGAAGATGCGCCCAACCTTAACAACGCTCTCTGAGGGGATTTGAGGTATGACACCG 60
QY 368 TCAAGCGGAGCTGAGGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
DB 61 TCAAGTCCGCGCTGAGAGCTGTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 428 TTCCGCGCGGAGCAACCTTCTCTGCGGCGGCGGAGCTGAGGCTGCGGCTGCGG 487
DB 121 TCGCGCGGAGGAGAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 488 GGGGAGCTGAGCAAGGCGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 547
DB 181 GCGGAGCTCAAGGAGGCGGAGCTTC-AGGGCGCAACACATCTCCGCGGAGCT 239
QY 548 CCAAGCTCG 556
DB 240 CCGGCGCTCG 248

```

RESULT 8

US-09-938-842A-2559  
Sequence 2559, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
APPLICANT: Krieps, Joel  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SRIPI300-3

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	9.48;	Score 127;	DB 9;	length 1011;
Best Local Similarity	56.28;	Pred. No. 3.2e-24;		
Matches 266; Conservative	0;	Mismatches 195;	Indels 12;	Gaps

```

Oy 217 GAGCAGAGATGGGGGCGCCTCTCTGCTCAGGCTCTTCTTCCACAGCTCTTCTGTTCAAGGC 27 6
Db 159 GATCCTAGAAATTCGACGCCATTAATTATTCGTCTTCACTTCCACGACTCTTTGTCCAAAGA 228
Oy 277 TCGCAGCGATGCATCTTCTTCGACCCGGA-----GGGAGAAAGACCGCCGG 324
Db 229 TGTGATGATCGGTGTTCCTTAGACAGACAGAAACTCTACAGGGAGAGAAAGCTTCT 288
Oy 325 CCGAACCCTGAACCTGGGTGCGCGGCTTTGAGGTCACTGCACACACATCAAGCCGGAACGTGAG 384
Db 289 CCCAACATTAATTAATGAAGATACAAAATTTGTGCACAGAAATCAAGAACATTAATGAA 348
Oy 385 GCCGCGTCCCCGCGGTGTGCTGCGACATCCGTGCGTTCGCGCGCGCAGCGA 444
Db 349 TCCGATGTCTCGTGAGGTTGTTTCAATGGCGTATCTTTCACAAATTTGGTGTAGAGATGCT 408
Oy 445 ACCAACCTTCTCGCGGGGCCCACTCGAGCGTGGCCGTGGGGCGGGGACTCGACAGC 504
Db 409 ACAATCTCGGTGGGGGCGCTTACTTGAGATGTTCTGTGGGAAATAAAGATTCAAAACC 468
Oy 505 GCCACGCGCTGCTGCTGCACACAGAACCCCGCCGCCGACGCGCAGCTCTGCGCAGCTC 564
Db 469 GCACACTACAGAGCTTGCACACAACAACTTCCACTCCAGAGAGGGTTTAATCAGATC 528
Oy 565 ATCTCCCTGTTGGCAGGCGGCGCTGTGCGCGCGCATGACGCGCTGTGCGGCGG 624
Db 529 ATTGCTAGTTCTTAATTCAGAGCTCTCTCGGTTGAAACATGCTGCGCTTATAGGACGC 588
Oy 625 CACACATCGGAGGAGCGCGGCTGCACACATCTTTCGGGCGCGCATCTACGGGA 677
Db 589 CACACGATCGGAAAAGCACAATGTGCGACACTTCGATCCGGAATTTATGGAGA 641

```

RESULT 9  
US-09-923-876-3446

```

: GENERAL INFORMATION:
: APPLICANT: Lalgudi, Raghunath V.
: APPLICANT: Kamigaki, Laura Y. (Ito)
: APPLICANT: Sherman, Bradley K.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLINGS
: FILE REFERENCE: PL-0012-1 CON
: CURRENT APPLICATION NUMBER: US/09/923,876
: CURRENT FILING DATE: 2001-08-06
: PRIOR APPLICATION NUMBER: 09/298,329
: PRIOR FILING DATE: 1999-04-21
: PRIOR APPLICATION NUMBER: 60/085,331
: PRIOR FILING DATE: 1998-05-05
: NUMBER OF SEQ ID NOS: 6332
: SOFTWARE: PERL Program
: SEQ ID NO 3446
: LENGTH: 234
: TYPE: DNA
:
:
:

```

```

; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700162281H1
US-09-923-876-3446

```

Query Match	8.7%	Score 118.2	DB 10	Length 234
Best Local Similarity	71.2%	Pred No 3.7e-22		
Matches 156	Conservative 0	Mismatches 63	Indels 0	Gaps 0

[illegible]

RESULT 10  
US-09-938-842A-2693  
; Sequence 2693, Application US/09938842A

```

/ GENERAL INFORMATION:
/ APPLICANT: Harper, Jeff
/ APPLICANT: Kreps, Joel
/ APPLICANT: Wang, Xun
/ APPLICANT: Zhu, Tong
/ TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
/ TITLE OF INVENTION: SAME, AND METHODS OF USE
/ FILE REFERENCE: SCRIPI300-3
/ CURRENT APPLICATION NUMBER: US/09/938,842A
/ CURRENT FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: US 60/227,866
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: US 60/264,647
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/300,111
/ PRIOR FILING DATE: 2001-06-22
/ NUMBER OF SEQ ID NOS: 5379
/ SEQ ID NO 2693
/ LENGTH: 993
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
/
US-09-938-842A-2693

```

Query Match	8.7%	Score 117.4;	DB 9;	Length 993;
Best Local Similarity	48.5%;	Pred. No. 1.2e-21;		
Matches 439; Conservative	0;	Mismatches 436;	Indels 30;	Gaps 3

QY	128	ACGCAAGCTCTCGCCCAAGCTTCTATGGGTCTCTCTGCCCCAACCTGAGAGCAATGCTTC	187
Db	86	ATGCTCAACTGTGCAGAGATTACTAGCGCTCGACATGTCTCTAGGGTAGAGCTCATCGTTA	144
QY	188	GGCGCGCATGACCCAGAGCCGTCGCACATGTAGCAGAGAGATGGCGCCTCTGCTCAAGCC	247
Db	146	ACGACGCGGTTACTACAAATTCAAAACAACATGTACAAAGGGCTCTGCACAGGTGGCGGA	205
QY	248	TCTTCTTCACAGCTGCTTTCGTTCAAGCCTCGACGATCGATCTTCTC-----	297
Db	206	TCTTCTTTCACAGCTGCTTCGTCGAGGAGTGTGATCGCTGTGTATTATACATCTGAGA	265
QY	298	--GAGCGCGAGAGGAGAGAAGACCGCGGAGCGCAACTCAATCTGGTGGCGGCTTTAGG	355
Db	266	ATGAAGACGCAAGAAAGCAGATGACAAATTAATCTCTGCGCCGAGACGAGATTGGACA	322

[illegible]







```

OY 380 TCGAGGCGCGCTGCCCCGCGGCGTGTGTCGTGCGCCGACATCCGCGGCTTGCGGCGCG 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 314 TAGAAGCTCATGCTTCCACAGTCTCATGTGTGTCATGTGTCATGTCATGTCACACAGTG 373
OY 440 ACGGAACCACTTCTCGCGCGCGCGACCTGAGACCTGCGGCTCGGCGCGGACCTGA 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 ACTCGGTGCTTACCGCGAGCCCAAGCTACAGCATCCCGACGAGGAGCGTGA 433
OY 500 CGAGCGCAGCGCTCGCTCGCCCAACAGAACCCCGCGCGCGCGCGCGCGCGCGCG 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 GGGTCTCAAC-----AATCTGATGTAACCTTACCGGGTCCACGATCTCGCTCTG 487
OY 560 CGCTATCTCCCTGTTTCGCGAGGAGCGCTGTGCGCGCGCGCGCGCGCGCGCGCG 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 488 GAGCCTGAGTTATTCAGAACAGAGGATGAACAGCTTCATGTCATGCACTCTTTGG 547
OY 620 GCGCGACACATCGGCGAGGCGCGGTGACACACCTTCGCGCGCGCGCGCGCGCGCA 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 548 GTGCAACACACTGTGTGTCAGAAATGTGTCTCTTATGTCAGAAATCACTACCTTCC 607
OY 680 CCGACA-----TCACGCTCTCTTGTGCGGCGCTGCGCGCGCGCGCGCGCGGT 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 AAGAACTGGAGCAGCGGACCGCTGCANAGACCCCGCTTGTGTACACCTTAAGAAACA 667
OY 731 CCGCGCGCGACGCGCAACTGCGCGCCCATTCAGCTGCAGACCGCGGTGAGTTGACAGG 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 CATGCAAAATAGCGGAGCGGCGCTAGACAGTGCAGTGCATTTGAGATTCGACACAC 727
OY 791 CCTACTTCAACAACTGTGTGCGCGCGCGCGCTGTTCACATGCGACGACGAGACTTCA 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 AGTCTTCAAGCAAAATCGTAAAGAGAGAGAGTGTGCAAGTTGACCAACCGCTCGCAT 787
OY 851 ACGCGGCTGCGAGGACCGGTGTGAGGAGAGTACAGCGCGCGCGCGCGCGCTTCAAG 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 CCGACCCACAGACTGTGTGAGATGTGTGCTGTGTGTATTAACAGCGCTTCTTCAACG 847
OY 911 CCGACTTGTGCGACCTGATTTAGATGGGCAACGTTGGGTGTCACCGCGCACCGCG 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 848 GTGAGTCTGTAGCAAAATGTTGAAGATGGAGAGATGTGATGTGCTTACTGTGTGTAAG 907
OY 971 GACAGATCAGCGCAACTGCGGCTGCTCAAC 1002
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 908 GTGAGATCAGAAAGAACTGCAGAGAGATTCAC 939

```

## RESULT 14

```

US-09-938-842A-783
; Sequence 783, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 783
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-783

```

```

Query Match      8.4%; Score 113.4; DB 9; Length 981;
Best Local Similarity 51.7%; Pred. No. 1.4e-20;
Matches 286; Conservative 0; Mismatches 261; Indels 6; Gaps 1;

OY 126 CCACGACAGCTCTCGGCCACGTTCTATGCGTCTCTGCCCCCAACCTGCAGAGACATCGT 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 CCACGCTCAATTCAGATGATGATTTCTATGCTCAATCTGTGCTTAATGCGAAAGATGT 125
OY 186 TCGGCGCGCGATGACCCAGCGCGCTGCAAGTACAGAGAGATGGGCGCTCTGCTCAG 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 TCAGATTTTGTTCAGAACACAGTTTCTATATGCTCTCTCTCTGCTGCTCATAG 185
OY 246 GCTCTTCTCCACAGCTGCTTCTTCAAGCTGCGACGATGATCTTCTGACGCG-- 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 AATGCAATTCATACATGTTTGTGCGAGGTGTGATGATCAGTGTATTAACCTCAAC 245
OY 304 ----GAGGGGAGAAAGCCGCGGCGGAACTGAACCTGAGTGTGCGCGCTTGAAGTCAAT 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GTCAAGAAACGACAGAGAGACGACTCTTAACCTAACGCTTCAGAGGTTTGCGCTTCAAT 305
OY 360 CGACACCATTCAGAGCGAAAGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGACAT 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 CGACCATTTAATCTGTGCTTGAAGCTCAATGCTGGAATTTGCTCTTCTGCTGATAT 365
OY 420 CTTGCGCTTCCGCGCGCGCGCGCAACCACTTCTCGCGCGCGCGCGCGCGCGCGCG 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 TATGCTCTTACTCTGCGACGCTGTCTGTTTCAACGCGAGGACCGAAATTTGAGTTCAC 425
OY 480 GCTCGGCGCGGAGCTGACGACGCGCGCGCGCGCTGCTGCGCAACACCCCGCGC 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 GACCGGAGAAAGACGCGGAGATATCAACGAGCGGAGGATTAAGCACATTTCTCC 485
OY 540 CCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 TCCAAACGATATATCAACCAATCTTCAGACACTTTTGCAACAGAGACTGTATTTAA 545
OY 600 CGACATGACGCGCTGTGCGCGCGCGCACACCACTCGGGCAGCGCGGTGACCACTTCG 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 GACCTCTTTTACTATTCGCGGCTCACACTATTTGTGTATCTACTCTCTCTTTCAC 605
OY 660 GCGCGCGATTCAC 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 AATCGCTCTTAC 618

```

## RESULT 15

```

US-09-938-842A-1518
; Sequence 1518, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1518
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1518
Query Match      8.1%; Score 110; DB 9; Length 987;

```

Best Local Similarity 49.0%; Pred. No. 1,2e-19;  
Matches 429; Conservative 0; Mismatches 420; Indels 27; Gaps 4;

```
OY 147 GTTCTATGCTCTCTCTCCCAACCTCCAGACATCGTTGGGGGCGGATGACCGAGGC 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 GTTTTACTTAACTACGTCTGCTAGAGCTGAACTATGTCCGAACGCCGTGAATGCTGG 173
OY 207 CGTCGCAAGTACAGAGATGGGCGCTCTCTGCTAGGCTCTCTCTCCAGACTGCTT 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 TTTCAGTTGACCCAGAAATGCAACCGGAATGATGATGATTCACAGACTGTTT 233
OY 267 CGTTCAGGCTGAGAGATGATCTCTCT--CGACCGGAGGAGAGAACCGCGCG 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 TGTCCAAAGTTGTGACGTTGATCTTATTCGAGACTAACCCGAGAGAACCGCGG 293
OY 324 GCCGAACTGAACTGCGGCTTTGAGTCAATGACACATCAAGCGGAAGCTGGA 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 TCCGAACCTCA--TCTCAAGGATTTGAATCATAGAACGCAAAACGACGCTGA 350
OY 384 GGGCGGTGCGCGCGCTGCTGCGCGGACATCCCGCTTGCAGCGCGGACG 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 AGCCGCTGCTGAGAGTGTCTTGTGCTGATATTTTAGCANTAGCCGCTGTGACAC 410
OY 444 AACCACCTTCTGCGCGGCGCACTGAGAGGTGCGCTGCGGCGGCGGACTGAGAC 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 AGTATTCTCACTAAGGAACAGGCTGGAGATACCAAGGAGCTAGAGATGTTAGT 470
OY 504 GGGCAGCGCTCGCTCGCCACAGCAACCCCGCGCCGACGCGCAGCTCGGACGCT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 TTCTTGGCTTCGAATGTAA--CAATCTCCCTGCTGCCGTGATCTCCCTTCA 527
OY 564 CATCTCCCTGTTGGGAGGAGGCGCTGCGCGCGGACATGACGCGGCTGTGCGGCGC 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 528 ACAACAGAAATTTCTCGCCTTGGGACTCAATACCTGTGATCTCTCTCTCGTGGAG 587
OY 624 GCACACATCGCGGAGCGCCGCTGACACACTTCGCGGCGCGCATCTACGGGACACG 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 588 ACACAGCATCGGAACAGAGGATGTGTATTCAGAAACAGGCTATTCAATACACCGG 647
OY 683 -----ACATCAAGCCTCTTCCGCGGCGCTGCGGCGAGCAGAGTGGCC 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 ACAAAACGAGATGCAACCATGCAACCAATTTTGGCGGAGCTTCAAAACAAATGTCC 707
OY 726 GCGGTCCGGGCGGAGGAGCACTGCGCCCATGACGTGACAGCGCGGTGAGTTGA 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 708 CCAAAAGCGGAGGTTCACTGCGGCTGATCTGACACCGGAAGCGGAAGCACTTGGGA 767
OY 786 CACGCGCTACTTCAACCAACTGCTGTGCGGCGGCGCTGTTCACCTGAGACGAGGCT 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 768 CACTTCTTACTACAACTAAGCCTAGCCGTGGCGGTGAGTCTCCAAATCGATCAAGTCT 827
OY 846 CTTCAAGCGGCGGTCGAGAGCGGCTGCTGAGGCAATACAGCGCCTCGCTCTT 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 828 TTGACACGATTCACGCCACTGACCCATAGACCATATGATGCTCTAGAGACACTT 887
OY 906 CAAGCGGACTTCTGCTGAGCAAGCATGATTAGGCAAGCTTGGGCTGCTACCGGAC 965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 888 CAAGCTTGAATTGCGAGGTCAATGGTCAAGATGATATCGGTGTGTTACCGAGC 947
OY 966 CGCGGACAGATCAAGCGCAACTGCGGCTGCTCAA 1001
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 948 TAATGGGAAATCCGTAGGCTTGTCTCTCGGTTAA 983
```

Search completed: May 3, 2003, 10:14:37  
Job time : 137 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:32:18 ; Search time 30 Seconds  
(without alignments)  
306,979 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 1604  
Sequence: 1 MASPTLMQGLVAVSLSCVA.....VGLTGTAQIRRCRVVNS 313

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	965	60.2	318	US-09-615-192A-389	Sequence 389, App
2	950	59.2	323	US-09-615-192A-395	Sequence 395, App
3	899.5	56.1	315	US-09-615-192A-378	Sequence 378, App
4	846	52.7	266	US-09-615-192A-394	Sequence 394, App
5	822	51.2	333	US-09-615-192A-394	Sequence 2, App1
6	807	50.3	287	US-09-615-192A-387	Sequence 387, App
7	714.5	44.3	324	US-09-365-150-2	Sequence 32, App1
8	664	41.4	351	US-09-615-192A-397	Sequence 397, App
9	624	38.9	201	US-09-615-192A-391	Sequence 391, App
10	605.5	37.7	324	US-08-671-320-13	Sequence 13, App1
11	605.5	37.7	324	US-08-668-577-13	Sequence 13, App1
12	604.5	37.7	324	US-08-671-320-11	Sequence 11, App1
13	604.5	37.7	324	US-08-668-577-11	Sequence 11, App1
14	585.5	36.5	349	US-08-190-029A-10	Sequence 10, App1
15	585.5	36.5	349	US-08-462-695-10	Sequence 10, App1
16	549	34.2	313	US-08-671-320-15	Sequence 15, App1
17	549	34.2	313	US-08-668-577-15	Sequence 15, App1
18	543	33.9	313	US-08-671-320-17	Sequence 17, App1
19	543	33.9	313	US-08-668-577-17	Sequence 17, App1
20	536	33.4	161	US-09-615-192A-388	Sequence 388, App
21	530.5	33.1	308	US-09-615-192A-381	Sequence 381, App
22	528.5	32.9	223	US-09-615-192A-385	Sequence 385, App
23	476	29.7	208	US-09-615-192A-401	Sequence 401, App
24	406	25.3	143	US-09-615-192A-332	Sequence 332, App
25	365.5	22.8	170	US-09-615-192A-399	Sequence 399, App
26	348	21.7	157	US-09-615-192A-376	Sequence 376, App
27	339.5	21.2	179	US-09-615-192A-376	Sequence 376, App

28	333.5	20.8	202	US-09-615-192A-386	Sequence 386, App
29	320	20.0	120	US-09-615-192A-392	Sequence 392, App
30	305	19.0	95	US-09-615-192A-390	Sequence 390, App
31	294.5	18.4	138	US-09-615-192A-384	Sequence 384, App
32	291	18.1	120	US-09-615-192A-393	Sequence 393, App
33	289	18.0	111	US-09-615-192A-379	Sequence 379, App
34	252.5	15.7	118	US-09-615-192A-333	Sequence 333, App
35	235	14.7	114	US-09-615-192A-347	Sequence 347, App
36	221	13.8	103	US-09-615-192A-398	Sequence 398, App
37	198.5	12.4	117	US-09-615-192A-400	Sequence 400, App
38	176.5	11.0	121	US-08-190-029A-12	Sequence 12, App1
39	176.5	11.0	121	US-08-462-695-12	Sequence 12, App1
40	132.5	8.3	362	US-09-596-824-6	Sequence 6, App1
41	119.5	7.5	389	US-09-596-824-2	Sequence 2, App1
42	112	7.0	366	US-09-596-824-4	Sequence 4, App1
43	108.5	6.8	344	US-09-024-532-4	Sequence 4, App1
44	108.5	6.8	345	US-07-792-259-12	Sequence 12, App1
45	108.5	6.8	364	US-07-792-259-17	Sequence 17, App1

#### ALIGNMENTS

RESULT 1  
US-09-615-192A-389  
; Sequence 389, Application US/09615192A  
; Patent No. 6410718  
; GENERAL INFORMATION:  
; APPLICANT: Blockberg, Leonard N.  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1003C4U  
; CURRENT APPLICATION NUMBER: US/09/615,192A  
; CURRENT FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 08/975,316  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: US 08/713,000  
; PRIOR FILING DATE: 1996-09-11  
; PRIOR APPLICATION NUMBER: US 09/169,789  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 405  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 389  
; LENGTH: 318  
; TYPE: PR1  
; ORGANISM: Pinus radiata  
; US-09-615-192A-389

Query Match 60.2%; Score 965; DB 4; Length 318;  
Best Local Similarity 60.1%; Pred. No. 2,4e-95;  
Matches 191; Conservative 39; Mismatches 76; Indels 12; Gaps 2;  
Query 7 MCLVAVSLSC-----VAHAIOLSPFYASSCPNLOSIYAAATOVAASRORGASTL 58  
Db 1 MRLVLCGLAAVAVAFIHNAVAGQLSTFYASCRLRISYKVAQVAKAEKRGASL 60  
Query 59 LRLFFHDCFFVGGDGSITLDAG----GEKTAGPNLMSVGRFEVIDTIKRNVEACPGVVS 114  
Db 61 VRLHFFHDCFFVGGDGSITLDAG----GEKTAGPNLMSVGRFEVIDTIKRNVEACPGVVS 120  
Query 115 CADLITLAAARDGNLGGFTWSPPLGRDSTTASASLANSNPPPPASITGLTISLGRG 174  
Db 121 CADLITLAAARDGNLGGFTWSPPLGRDSTTASASLANSNPPPPASITGLTISLGRG 180  
Query 175 LSPDMTALSGAHTIGACTTFRGRTYGTDTINAFALROOTCPRSGGDMAPIDVO 234  
Db 181 LSTYDVALSGAHTIGACTTFRGRTYGTDTINAFALROOTCPRSGGDMAPIDVO 240  
Query 235 TPVAFDTATYTNLSSRGLFHSQDELFGNGSDALYRQISASLSLNADPVAAMITMGV 294  
Db 241 TSIFDKKYSNLKIQGLLHSDQDLFNGSGSTDSQVTAAYSNSNSPFIETAAVYKMGNI 300



Db 1 MASFAMSLAFIALMCSFVAAYALSATFENTSCPKLSTVOAAVQAANERKRGASL 60  
QY 59 LRLPHDFHDCVCGSGSILLDAG-----GEKTAGPMLNSVRGEVIDTIRKRVFAACGCVS 114  
Db 61 LRLPHDFHDCVCGSGSVLLDSSFTLGEKTAVPNNNSARGFVDITKSGVEAVCGSVS 120  
QY 115 CADILALARGDTNLLGGPTWSPVPLGRDSTTASASLANSPPEPTASLGTLISLFRG 174  
Db 121 CADILALARGDSTVQLGGPTWVQLGRDSTASLSCANNINIPAPTSNLSALISLFRG 180  
QY 175 LSPRMTLSGAHTIGARCTTFRGRITGDTDIASFAALROOTCPSSGSGDNLAPIDV 234  
Db 181 LSTKMVVALSAGHTIGARCTSFARATYNESNINAAVATSLKTNCPPTGSDNNLSPLDRV 240  
QY 235 TPVRFDYAFETNLSRGLFHSDOEL 260  
Db 241 TPTRFDINYSNLSQKGLHSDQOL 266

## RESULT 5

US-09-097-319A-2  
; Sequence 2, Application US/09097319A  
; Patent No. 6384207  
; GENERAL INFORMATION:  
; APPLICANT: Atinley, Michael  
; APPLICANT: Armstrong, Katherine  
; APPLICANT: Belmar, Scott  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Hopkins, Nicole  
; APPLICANT: Menke, Michael A.  
; APPLICANT: Paredy, Dayakar  
; APPLICANT: Petolino, Joseph F.  
; APPLICANT: Smith, Kelley  
; APPLICANT: Woosley, Aaron  
; TITLE OF INVENTION: Regulatory Sequences for Transgenic Plants  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dowelance Patent Department  
; STREET: 9330 Zionsville Road  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: USA  
; ZIP: 46268  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/097,319A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stuart, Donald R  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317 337 4816  
; TELEFAX: 317 337 4847  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 333 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-097-319A-2

Query Match 51.2%; Score 822; DB 4; Length 333;  
Best Local Similarity 51.1%; Pred. No. 5.8e-80;  
Matches 168; Conservative 51; Mismatches 90; Indels 20; Gaps 4;  
QY 5 TLMQCLVAVSLSCVAHQ-----LSPFTYASSCPNLSQSYRAAMTOAVSSEORHGA 56  
Db 3 TSMGLVILCLVSLSPASVILGHPWGILFPQFYDHSCPKAKETIVSIVAQAVAKETRMMA\*62

QY 57 SLRLPHDFHDCVCGSGSILLDAG-----GEKTAGPMLNSVRGEVIDTIRKRVFAACPGV 112  
Db 63 SLRLPHDFHDCVCGSGSVLLDSSFTLGEKTAVPNNNSARGFVDITKSGVEAVCGSVS 120  
QY 113 VSCADILALARGDTNLLGGPTWSPVPLGRDSTTASASLANSPPEPTASLGTLISLFRG 172  
Db 123 VSCADILALARGDSTVQLGGPTWVQLGRDSTASLSCANNINIPAPTSNLSALISLFRG 180  
QY 173 OGLSPRMTLSGAHTIGARCTTFRGRITGDTDIASFAALROOTCPSSGSGDNLAPIDV 234  
Db 183 OGLNVDVVALSGHTIGARCTTFRGRITGDTDIASFAALROOTCPSSGSGDNLAPIDV 240  
QY 236 GNLAPIDVOTPVREDYAFETNLSRGLFHSDOELFNGSGOD-ALVROYASASLFFNADF 284  
Db 243 NNLPLDFETTPAKFENFYKMLLAGKGLSDEILTKSATAALVAVAADVNLFFQHF 302  
QY 285 VAAITRMGNVGLTGTACIIRNCRVNS 313  
Db 303 AOSMVMGNISPLTIGSGEIRKNCRLNN 331

## RESULT 6

US-09-615-192A-387  
; Sequence 387, Application US/09615192A  
; Patent No. 6410718  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; APPLICANT: Havukkala, Ilkka  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1003c4U  
; CURRENT APPLICATION NUMBER: US/09/615,192A  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 08/975,316  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR FILING DATE: 1996-09-11  
; PRIOR APPLICATION NUMBER: US 09/169,789  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 405  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 387  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
; US-09-615-192A-387

Query Match 50.3%; Score 807; DB 4; Length 287;  
Best Local Similarity 57.4%; Pred. No. 1.9e-78;  
Matches 162; Conservative 35; Mismatches 75; Indels 10; Gaps 3;  
QY 9 CLVAVSLSC-----VAHQSLPTFYASSCPNLSQSYRAAMTOAVSSEORHGA 64  
Db 6 CLVAVSLSC-----VAHQSLPTFYASSCPNLSQSYRAAMTOAVSSEORHGA 64  
QY 65 DCFVCGSGSILLDAG-----GEKTAGPMLNSVRGEVIDTIRKRVFAACGCVS 120  
Db 66 DCFVCGSGSVLLDSSFTLGEKTAVPNNNSARGFVDITKSGVEAVCGSVS 120  
QY 121 LARGDTNLLGGPTWSPVPLGRDSTTASASLANSPPEPTASLGTLISLFRG 180  
Db 126 LARGDTNLLGGPTWSPVPLGRDSTTASASLANSPPEPTASLGTLISLFRG 180  
QY 181 TALSGAHTIGARCTTFRGRITGDTDIASFAALROOTCPSS- GSGDNLAPIDVOTPV 238  
Db 186 VALSGAHTIGARCTTFRGRITGDTDIASFAALROOTCPSS- GSGDNLAPIDVOTPV 238  
QY 239 PFTAFETNLSRGLFHSDOELFNGSGODALVROYASASLFF 280  
Db 246 PDKNTYINLKSRLHSDQELFNGSGODSHVYTYASQNTF 287

RESULT 7  
US-09-365-150-2  
Sequence 2, Application US/09365150  
Patent No. 6278041  
GENERAL INFORMATION:  
APPLICANT: Lagimod, Mark  
TITLE OF INVENTION: No. 6278041el Peroxidase Gene Sequences  
FILE REFERENCE: S-31081P1  
CURRENT APPLICATION NUMBER: US/09/365,150  
CURRENT FILING DATE: 1999-07-30  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 2  
LENGTH: 324  
TYPE: PRF  
ORGANISM: Nicotiana tomentosiformis  
US-09-365-150-2

Query Match 44.5%; Score 714.5; DB 4; Length 324;  
Best Local Similarity 47.2%; Pred. No. 1.9e-68;  
Matches 143; Conservative 53; Mismatches 97; Indels 17; Gaps 6;

OY 10 LVAVSLSCVAHAOLSPFYASSCPNLSQIVRAAMQAVASEQRMGASLRLFFHDCFEVQ 69  
DB 13 LVAFMAAS--NAQLSATFYDSTCPNTSIVRGVMDQRORDARAGAKIRLHFHDCFEV 69  
OY 70 GCDGSILLDAGC---EKTAGPNLNSVRGEVYDITKRNVEACGCVVSCADILALARDG 126  
DB 70 GCDGSILLDITDGTOTENKAAPNVGA-GGFDIVDITKALEVCGVVCADILSLASEIG 128  
OY 127 TNLGGPTWSVPLGRDSTASLANSNPPPTASLGTLSLFGROGLSPRDMTALSGA 186  
DB 129 VALAGPFWQVLFGRKNSLITNRSEANSIDISPEFTRAVMPLFTNNGMDITDLVAQSGA 188  
OY 187 HTIGARCTEGRITG-----DTIDNASFAALRQOTCPRSSGCGDN-LAPIDVOTPV 238  
DB 189 HTFGARCTEGRITG-----DTIDNASFAALRQOTCPRSSGCGDN-LAPIDVOTPV 248  
OY 239 EDTAFVTLNLRGLFHSDDLEF--NGSSODALVROYASASLNFADPVAAMIRMGV 296  
DB 249 FDNDFVTLNLRGLFHSDDLEF--NGSSODALVROYASASLNFADPVAAMIRMGV 308  
OY 297 LTGTAGQIRRCRVN 312  
DB 309 LTGTAGQIRRCRVN 324

RESULT 8  
US-09-615-192A-397  
Sequence 397, Application US/09615192A  
Patent No. 6410718  
GENERAL INFORMATION:  
APPLICANT: Bloksberg, Leonard N.  
TITLE OF INVENTION: Materials and Methods for the  
FILE REFERENCE: 11000.1003c4U  
CURRENT APPLICATION NUMBER: US/09/615,192A  
CURRENT FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 08/975,316  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11  
PRIOR APPLICATION NUMBER: US 09/169,789  
PRIOR FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 405  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 397  
LENGTH: 351  
TYPE: PRF  
ORGANISM: Pinus radiata  
US-09-615-192A-397

Query Match 41.4%; Score 664; DB 4; Length 351;  
Best Local Similarity 47.0%; Pred. No. 5.7e-63;  
Matches 142; Conservative 55; Mismatches 93; Indels 12; Gaps 5;

OY 22 AQLSPFYASSCPNLSQIVRAAMQAVASEQRMGASLRLFFHDCFEVQCGGSLDLD-G 80  
DB 31 AQLSPFYASSCPNLSQIVRAAMQAVASEQRMGASLRLFFHDCFEVQCGGSLDLD-G 90  
OY 81 GKTAGPNLN-SVRGEVYDITKRNVEACGCVVSCADILALARDGNTLGGPTWSVPL 139  
DB 91 GKTAGPNLN-SVRGEVYDITKRNVEACGCVVSCADILALARDGNTLGGPTWSVPL 150  
OY 140 GRDSTT-ASASLANSNPPPTASLGTLSLFGROGLSPRDMTALSGAHTIGARCTFR 198  
DB 151 GRDSTT-ASASLANSNPPPTASLGTLSLFGROGLSPRDMTALSGAHTIGARCTFR 210  
OY 199 GRIG-----DTIDNASFAALRQOTCPRSSGCGDN-LAPIDVOTPVDPTRAVFNLSR 250  
DB 211 NRYNSTGTQMRDPTMDSFAKNLYLTCPTS-TTVNTTKDIRPNVFDNKYYDILNR 269  
OY 251 RGLFHSDDLEFNGSSODALVROYASASLNFADPVAAMIRMGVLTGTAGQIRRCRV 310  
DB 270 QITFTSDQTLTYDTRIDIVINFAVQSLFFEQVLSMLKMGQDLVLSGSEGRKNCMA 329  
OY 311 VN 312  
DB 330 AN 331

RESULT 9  
US-09-615-192A-391  
Sequence 391, Application US/09615192A  
Patent No. 6410718  
GENERAL INFORMATION:  
APPLICANT: Bloksberg, Leonard N.  
TITLE OF INVENTION: Materials and Methods for the  
FILE REFERENCE: 11000.1003c4U  
CURRENT APPLICATION NUMBER: US/09/615,192A  
CURRENT FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 08/975,316  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11  
PRIOR APPLICATION NUMBER: US 09/169,789  
PRIOR FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 405  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 391  
LENGTH: 201  
TYPE: PRF  
ORGANISM: Pinus radiata  
US-09-615-192A-391

Query Match 38.9%; Score 624; DB 4; Length 201;  
Best Local Similarity 64.7%; Pred. No. 4.8e-59;  
Matches 130; Conservative 21; Mismatches 44; Indels 6; Gaps 2;

OY 1 MASPTLQCLVAVSL--SCVAHAOLSPFYASSCPNLSQIVRAAMQAVASEQRMGASL 58  
DB 1 MASPTLQCLVAVSL--SCVAHAOLSPFYASSCPNLSQIVRAAMQAVASEQRMGASL 60  
OY 59 LRLFFHDCFEVQCGGSLDLDAG-----EKTAGPNLNSVRGEVYDITKRNVEACGCVV 114  
DB 61 LRLFFHDCFEVQCGGSLDLDAG-----EKTAGPNLNSVRGEVYDITKRNVEACGCVV 120  
OY 115 CADILALARDGNTLGGPTWSVPLGRDSTTASASLANSNPPPTASLGTLSLFGROG 174  
DB 121 CADILALARDGNTLGGPTWSVPLGRDSTTASASLANSNPPPTASLGTLSLFGROG 180  
OY 175 LSPRDMTALSGAHTIGARCT 195

Db 181 LSTKMTIVLSGARTIGARCT 201

## RESULT 10

US-08-671-320-13  
; Sequence 13, Application US/08671320  
; Patent No. 5840558  
; GENERAL INFORMATION:  
; APPLICANT: VIERLING JR, RICHARD A  
; TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
; TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: 1125 SO. 103RD STREET  
; STREET: SUITE 330  
; CITY: OMAHA  
; STATE: NE  
; COUNTRY: US  
; ZIP: 68124-1076  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/671,320  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JONDLER, ROBERT J.  
; REGISTRATION NUMBER: 33,915  
; REFERENCE/DOCKET NUMBER: 1227-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 402-398-9000  
; TELEFAX: 402-398-9005  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 324 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-671-320-13

Query Match 37.7%; Score 605.5; DB 2; Length 324;  
Best Local Similarity 42.5%; Pred. No. 9.7e-57;  
Matches 137; Conservative 54; Mismatches 112; Indels 19; Gaps 9;

QY 6 LMOCIVAVSLSCVAHAOLSPFTFVASSCPNLSQIVRAAMTOAVASEORMGASLIRLFPHD 65  
DB 8 LSTICLLA--LASTHAOLQIGFYAKSCPNAEQIVLKFVHDHINAPSLAALIRMHFD 64  
QY 66 CFVGGGDSIILDA--GGEKTAGPNLNSVGFVEYIDTIKKNVENACPGVASCADILALA 122  
DB 65 CFVRGCDASVLNSTNOAEKNAPNL-TVRGFEDIRIKSLVERECGCVVSCADILALS 123  
QY 123 ARGTNLLGPTWVPLGRDSTTASIASLANSNPPPTASIGTLISLFGROGLSPDMTA 182  
DB 124 ARDTIVATGPPWKVPTGRDGVISNLTPEARINIPAPSNFTTLOTLEFANGGLDKDLVL 183  
QY 183 LSGAHTIGARCTTFRGRIV-----GDPD--INASFAA-LROQTCPRSGGNGNLAIPDYO 234  
DB 184 LSGAHTIGAHSSLSNRLEFNTGKDDPSIDSEYANLAKFKC-TDLNLTNTKTIEMD 242  
QY 235 TPVR--FDYAFYTNLISRGFLHSDQELF-NGGSODALVROYASASISFNADFYAAMIRM 291  
DB 243 PGRKRTFDLSYSHVTKRKGLESDAALLTNSVTAKQIITELLEGSVENFPAEFATSMEM 302  
QY 292 GNVGVLGTAGQIRNRCRVNS 313  
DB 303 GRINVTGTGEGIRKHCALNS 324

RESULT 11  
US-08-868-577-13  
; Sequence 13, Application US/08868577  
; Patent No. 5866695  
; GENERAL INFORMATION:  
; APPLICANT: Vierling Jr., Richard A  
; TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
; TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz  
; STREET: 555 13th Street NW, Suite 701 East  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/868,577  
; FILING DATE: 04-JUN-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jondler, Robert J.  
; REGISTRATION NUMBER: 33,915  
; REFERENCE/DOCKET NUMBER: N1227-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 402-333-1550  
; TELEFAX: 402-333-1510  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 324 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-868-577-13

Query Match 37.7%; Score 605.5; DB 2; Length 324;  
Best Local Similarity 42.5%; Pred. No. 9.7e-57;  
Matches 137; Conservative 54; Mismatches 112; Indels 19; Gaps 9;

QY 6 LMOCIVAVSLSCVAHAOLSPFTFVASSCPNLSQIVRAAMTOAVASEORMGASLIRLFPHD 65  
DB 8 LSTICLLA--LASTHAOLQIGFYAKSCPNAEQIVLKFVHDHINAPSLAALIRMHFD 64  
QY 66 CFVGGGDSIILDA--GGEKTAGPNLNSVGFVEYIDTIKKNVENACPGVASCADILALA 122  
DB 65 CFVRGCDASVLNSTNOAEKNAPNL-TVRGFEDIRIKSLVERECGCVVSCADILALS 123  
QY 123 ARGTNLLGPTWVPLGRDSTTASIASLANSNPPPTASIGTLISLFGROGLSPDMTA 182  
DB 124 ARDTIVATGPPWKVPTGRDGVISNLTPEARINIPAPSNFTTLOTLEFANGGLDKDLVL 183  
QY 183 LSGAHTIGARCTTFRGRIV-----GDPD--INASFAA-LROQTCPRSGGNGNLAIPDYO 234  
DB 184 LSGAHTIGAHSSLSNRLEFNTGKDDPSIDSEYANLAKFKC-TDLNLTNTKTIEMD 242  
QY 235 TPVR--FDYAFYTNLISRGFLHSDQELF-NGGSODALVROYASASISFNADFYAAMIRM 291  
DB 243 PGRKRTFDLSYSHVTKRKGLESDAALLTNSVTAKQIITELLEGSVENFPAEFATSMEM 302  
QY 292 GNVGVLGTAGQIRNRCRVNS 313  
DB 303 GRINVTGTGEGIRKHCALNS 324

RESULT 12  
US-08-671-320-11  
; Sequence 11, Application US/08671320  
; Patent No. 5840558

GENERAL INFORMATION:  
APPLICANT: VIERLING JR, RICHARD A  
TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: 1125 SO. 103RD STREET  
STREET: SUITE 330  
CITY: OMAHA  
STATE: NE  
COUNTRY: US  
ZIP: 68124-1076  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/671,320  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: JONDLE, ROBERT J.  
REGISTRATION NUMBER: 33,915  
REFERENCE/DOCKET NUMBER: 1227-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 402-398-9005  
TELEFAX: 402-398-9005  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 324 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-671-320-11

Query Match 37.7%; Score 604.5; DB 2; Length 324;  
Best local similarity 42.2%; Pred. No. 1.2e-56;  
Matches 136; Conservative 57; Mismatches 110; Indels 19; Gaps 9;

QY 6 LMOCIVAVSLSCVAHAQSLPTFYASCPNLSQSTVRAAMTOAVASEORMGASLRLFFHD 65  
DB 8 LSLCLIA---LIASTHAQLOLGFYANSCPKAEQIVLKFVHDHINAPSLAAILRMHFD 64

QY 66 CFVCGDGSILIDA---GGEKTAGPNLNSVGFVEIDITIKRNVACPGVSCADILALA 122  
DB 65 CFVRGCDASVILNSTTNQAEKNAPPNL-TVRCGFDFIDRIKSLVBAECPGVSCADILTLA 123

QY 123 ARDGNILGPTWVPLGRDSTTASASLANSNPPPTASLGLTILSLFRGGLSPRDMTA 182  
DB 124 ARDTIVATGPFVKVPTGRDGVSNLTERRNNIPAPSSNFTTLOTLEFANQGLDKDLVL 183

QY 183 LSGAHTTIGARCTTFRGRLY-----GDTD--INASFA-LROOTCPRSGGDNLAIPIDVQ 234  
DB 184 LSGAHTTIGAHCSLSNRLFNFTGKGDOPSLDSEYANLAKFKC-TDLNKLNTTKIEMD 242

QY 235 TPVR--FDPAVFTNLSRGLFHSDOELF-NGSODALVROYASASLNFADFAAMITM 291  
DB 243 PGRKTFDLSTYSHYIKRGLFESDAALLTNSVTKAQITQLLEGSEVENFAEFATSTIEKM 302

QY 292 GNVGVLTGTAGQIRNCRVNS 313  
DB 303 GRINVKTGEGEIRKHCAPINS 324

RESULT 13  
US-08-868-577-11  
Sequence 11, Application US/08868577  
Patent No. 5866695  
GENERAL INFORMATION:  
APPLICANT: Vierling Jr., Richard A  
TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY

NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz  
STREET: 555 13th Street NW, Suite 701 East  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/868,577  
FILING DATE: 04-JUN-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Jondle, Robert J.  
REGISTRATION NUMBER: 33,915  
REFERENCE/DOCKET NUMBER: N1227-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 402-333-1550  
TELEFAX: 402-333-1510  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 324 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-868-577-11

Query Match 37.7%; Score 604.5; DB 2; Length 324;  
Best local similarity 42.2%; Pred. No. 1.2e-56;  
Matches 136; Conservative 57; Mismatches 110; Indels 19; Gaps 9;

QY 6 LMOCIVAVSLSCVAHAQSLPTFYASCPNLSQSTVRAAMTOAVASEORMGASLRLFFHD 65  
DB 8 LSLCLIA---LIASTHAQLOLGFYANSCPKAEQIVLKFVHDHINAPSLAAILRMHFD 64

QY 66 CFVCGDGSILIDA---GGEKTAGPNLNSVGFVEIDITIKRNVACPGVSCADILALA 122  
DB 65 CFVRGCDASVILNSTTNQAEKNAPPNL-TVRCGFDFIDRIKSLVBAECPGVSCADILTLA 123

QY 123 ARDGNILGPTWVPLGRDSTTASASLANSNPPPTASLGLTILSLFRGGLSPRDMTA 182  
DB 124 ARDTIVATGPFVKVPTGRDGVSNLTERRNNIPAPSSNFTTLOTLEFANQGLDKDLVL 183

QY 183 LSGAHTTIGARCTTFRGRLY-----GDTD--INASFA-LROOTCPRSGGDNLAIPIDVQ 234  
DB 184 LSGAHTTIGAHCSLSNRLFNFTGKGDOPSLDSEYANLAKFKC-TDLNKLNTTKIEMD 242

QY 235 TPVR--FDPAVFTNLSRGLFHSDOELF-NGSODALVROYASASLNFADFAAMITM 291  
DB 243 PGRKTFDLSTYSHYIKRGLFESDAALLTNSVTKAQITQLLEGSEVENFAEFATSTIEKM 302

QY 292 GNVGVLTGTAGQIRNCRVNS 313  
DB 303 GRINVKTGEGEIRKHCAPINS 324

RESULT 14  
US-08-190-029A-10  
Sequence 10, Application US/08190029A  
Patent No. 5736363  
GENERAL INFORMATION:  
APPLICANT: EDWARDS, Richard Mark  
APPLICANT: BAWDEN, Lindsey  
TITLE OF INVENTION: IGF-II ANALOGUES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ALLERRETTI & WITCOFF, LTD.  
STREET: 10 S. WACKER DRIVE, SUITE 3000



Wed-May 7 14:10:02 2003

us-10-047-825-4.rai

Page 7

```

CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190.029A
FILING DATE: 28-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCV/GB92/01389
FILING DATE: 27-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 920401.7
FILING DATE: 05-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9116325.3
FILING DATE: 29-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: JOHN J. McDONNELL
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 94,062
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1234
TELEFAX: 312-715-1234
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-190-029A-10

Query Match
Best Local Similarly 36.5%; Score 585.5; DB 1; Length 349;
Matches 128; Conservative 34; Mismatches 104; Indels 15; Gaps

OY 23 QLSPTFFASCCRNQDSIYRAAMQAVASEQRMKASLLRLEFFHDCVCGCGSITLD--- 78
Db 2 QLTFFPTDNCPCWASNIYRDTIVNELRSDPIASIIIRLHFHDCVCGCASITLDNTTS 61

OY 79 AGCKETAGPNLNSVSGFEVIDITKRWNEACPGVASCADILALAAARDGTMLNGCPWSP 138
Db 62 FRTEDAFNGNANSAGFEVIDIRKKAIVESAECQPRVSCADILTAOOSVTLACGSPWRVP 121

OY 139 LGRDSTATASIASNPNPPFTASLGITLISLFGROGLS-PRDMTALSGAHTIOARCTTF 197
Db 122 LGRDSTLQAFILAVANANLPAPFTTLFQLKDSFRVWGINSRSDLVALSQGTFFKQNCRI 181

OY 198 NGRTYG-----DIDINASPAALROOTCPRSSGGDGLAPIDVDPYVPTFAFTNLLSR 250
Db 182 MRLYINFSNTGJLDPDTLNTFYLTQTLNGCLGPLNGMLSLAVDFDLRFTPFIFNKYYVLEQ 241

OY 251 RGLFHSPODELF---NGSGSOALIVROYASASASLFPADPVAM 288
Db 242 KGLISQDELFSPPNATDTPIIVRSRANSQTFFNATVEAM 282

RESULT 15
US-08-462-695-10
; Sequence 10 Application US/08462695
; Patent No. 5854025
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Richard Mark
; APPLICANT: BARDEN, Lindsay
; TITLE OF INVENTION: IGF-II ANALOGUES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & ALLEGRETTI, LTD.

```

```

STREET: 10 S. WACKER DRIVE, SUITE 3000
CITY : CHICAGO
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,695
FILING DATE: 5-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/190,029
FILING DATE: 28-FEB-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/01369
FILING DATE: 27-JUL-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9202401.7
FILING DATE: 05-FEB-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9116325.3
FILING DATE: 29-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: JOHN J. McDONNELL
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 94,062-A
TELECOMMUNICATIONS INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
INFORMATION FOR SEO ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-695-10

Query Match          36.5%; Score 585.5; DB 2; Length 349;
Best local similarity 45.6%; Pied. No. 1.5e-54;
Matches 128; Conservative 34; Mismatches 104; Indels 15; Gaps 4;

QY      23 QLSPTVYASCPNIQOSTIVRAAMQAVASQRMGSLRLRFHCFYQGCDSSILDD---- 78
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       2 QLPRTFYDSCSNVNIRDTIVNELRSRPIAASIIIRLFHCFFVNGDSILDDMTTS 61
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      79 AGGEKTAGPNINSVRGFEVIDITIKRNVEACPGVVS CADIIIALAADGTNLGGPTWSVP 138
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       62 FRREKAFENFANSARGPVVIDRKAAVESACPRTVSCADILLTMAOQSVALTAGPSWRVP 121
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      139 LGRSDTYSASLANSNPPEPTSLGLTSLISLEGROGLS -PRDVTALSGAHITGOARCTYE 197
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       122 LGRSDLSIAFLDLANMANIPAPFTTLQLCLDSFRNVGNLRSSDIYALISGCHFGKKQCREFI 181
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      198 RGRIVY-----PDTIDASFALRROTCPRSGGDGMILARDIVDPVPRPDYVFPTNLSR 250
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       182 MDLVLYFNSTGIPPTINTYVLTQTUNGCLCPINGNSLSALYDFDLRTFTJEDNKYYVINEEQ 241
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      251 RGLFHSDDELFI---NGSGDALVROYSAASLFFNADVAM 288
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       242 KGLIGSDDELFSSRNATDITPLVRSYANSTOTFFNAIVEAM 282
        ||::|||:::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||

Search completed: April 26, 2003, 12:36:14
Job time : 31 secs

```





Db 246 NTTFMDLITPAKFDKRYVGLTNNLGFESDAILTNATWKAIVDSFVSEATWTKFA 305  
Qy 286 AAMIRMGVGLTGTAGGRRRCRYVN 312  
Db 306 KSMKMGQIEVLGTGTGGRRCRYVN 332

## RESULT 2

US-10-027-559-6  
; Sequence 6, Application US/10027559  
; Patent No. US20020144307A1  
; GENERAL INFORMATION:  
; APPLICANT: Simmons, Carl  
; TITLE OF INVENTION: Maize Defense-Inducible Genes and Their  
; FILE REFERENCE: 035718/239835  
; CURRENT APPLICATION NUMBER: US/10/027,559  
; PRIOR FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: 60/243,120  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-027-559-6

Query Match 11.4%; Score 182.5; DB 12; Length 250;

Best Local Similarity 26.7%; Pred. No. 6.9e-10;  
Matches 74; Conservative 36; Mismatches 86; Indels 81; Gaps 14;

Qy 46 QAVASEQRMGASLIRLFHDCVQCGDSIILDAG--GKTAGP-----NL 89  
Db 24 RALIAEKSCAPMLRLAMH-----SAGTEPVSSTGCGPGTWHQSELAHGA 70  
Qy 90 NSVGFVEIDTIKRNVEAACPGVSCADILALADGNTLGGPTWGVPRGRDSTTASA 149  
Db 71 NA--GLDIANVRLEPIKEEP--ILSYADFYQAGVAVNEVYGGEIFPHGRD----- 121  
Qy 150 SLANSNPP-----PTASLGT--LISLFGHQ--GLSPRDMTALSAGHTTGACRTTFRGRIT 202  
Db 122 --KPOPPGGRLPDATKSGDHLRQVFGKMGCLSHQDIVALSGHTIG--RCHK----- 170  
Qy 203 GDTINASFALRQCTCPRSSGCGMLAPIDVQTPVRFPTAYFTMLLS--RRGLFH--SDQ 258  
Db 171 -----ERSGEGAM---TTNPVFNSTFKELLSGDMGSLFQLPSDK 209  
Qy 259 ELFNQGSQDALVROYASASLFNADPVAAIMRMGAVG 295  
Db 210 ALLSDPVFRPLVEKYADEKAFEDDYKKAHLKSELG 246

## RESULT 3

US-09-870-501-1  
; Sequence 1, Application US/09870501  
; Patent No. US20020100081A1  
; GENERAL INFORMATION:  
; APPLICANT: Takabe, Tetsuko  
; TITLE OF INVENTION: Peroxisomal Ascorbate Peroxidase Gene Induced by High  
; TITLE OF INVENTION: Temperature Stress and a Transgenic Plant Exhibiting  
; TITLE OF INVENTION: Thermotolerance  
; FILE REFERENCE: 026350-053  
; CURRENT APPLICATION NUMBER: US/09/870,501  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: JP 2000-172850  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 291  
; TYPE: PRT

; ORGANISM: Hordeum vulgare L. Haruna-nijo  
US-09-870-501-1

Query Match 11.0%; Score 176.5; DB 10; Length 291;

Best Local Similarity 25.6%; Pred. No. 3.3e-09;  
Matches 72; Conservative 42; Mismatches 112; Indels 55; Gaps 10;

Qy 36 LOSIVRAANT-QAVASEQRMGASLIRLFHDCVQ-----GCDGSIILDAGGKTAGP 87  
Db 11 LROYDARRARAFALIASKCAPIMRLAMHDAGTYDVNTRGANGSIRYEE--EETHGS 68  
Qy 88 NLNSVGFVEIDTIKRNVEAACPGVSCADILALADGNTLGGPTWGVPRGRDSTTASA 147  
Db 69 NAGLKAIDLLEPIK-----AKHPKITADLHQLAGVAVNEVYGGPTVERIPGRDS--- 120  
Qy 148 SASLANSNPPPTASLGT--LISLFGRLSPRDMTALSAGHTTGACRTTFRGRITGDT 205  
Db 121 --SVCEGRGLDPARKGAPHLRDIETFRMGLDKDIVALSGHSLGKHA----- 166  
Qy 206 DINASFALRQCTCPRSSGCGMLAPIDVQTPVRFPTAYFTMLL--SRGLFH--SDQELF 261  
Db 167 -----PERSGEGAM---TRDPLKFDNSYFLELLKGESEGLKLPDPTKALL 209  
Qy 262 NGGSQDALVROYASASLFNADPVAAIMRMGAVGLGTAG 302  
Db 210 DDEFFRRYVELYAKDEYFEKDYAESHKKLSLSELTFTPRSSG 250

## RESULT 4

US-09-734-017A-54  
; Sequence 54, Application US/09734017A  
; Patent No. US20020142422A1  
; GENERAL INFORMATION:  
; APPLICANT: Lerchl, Jens  
; APPLICANT: Renz, Andreas  
; APPLICANT: Ehrhardt, Thomas  
; APPLICANT: Reindl, Andreas  
; APPLICANT: Clirpus, Petra  
; APPLICANT: Bischoff, Friedrich  
; APPLICANT: Frank, Markus  
; APPLICANT: Freund, Annette  
; APPLICANT: Duenwig, Elke  
; APPLICANT: Schmidt, Ralf-Michael  
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv  
; TITLE OF INVENTION: the  
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides an  
; TITLE OF INVENTION: nucleosides  
; FILE REFERENCE: BASF-NAE-1331-99-US  
; CURRENT APPLICATION NUMBER: US/09/734,017A  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 60/171,100  
; PRIOR FILING DATE: 1999-12-16  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect  
; SEQ ID NO 54  
; LENGTH: 176  
; TYPE: PRT  
; ORGANISM: Physcomitrella patens  
US-09-734-017A-54

Query Match 8.5%; Score 136; DB 10; Length 176;

Best Local Similarity 29.0%; Pred. No. 1.4e-05;  
Matches 53; Conservative 25; Mismatches 71; Indels 34; Gaps 9;

Qy 30 ASSCPNLSIVRAAMTQA-----VASORNGASLIRLFHDCVQCGDSIILDAGGK 83  
Db 2 AKSYPNVSEKVALIEKARKIRGVAAEKNCAPITLRLAMH-----GSGTYDD--ESK 52  
Qy 84 TAGPNINSVR-----GFEVIDTIKRNVEAACPGVSCADILALADGNTLGG 132  
Db 53 TGGP-LGITIRFGQELAHGMGLDIAVNLLOPIKEQFP--ELSTVADFTYLAGVAVNEVYGG 110  
Qy 133 PTWSVPLGRDRSTTASASLANSNPPPTASLGTLLISLFGHQ--GLSPRDMTALSAGHTTIGQ 191

Db 111 PTFPHGRDHCPC---VEGRLPATKGLDHLRCVFTKQKMLTDRDIYVLSGATLIG- 166  
 QY 192 ARC 194  
 Db 167 -RC 168

## RESULT 5

US-09-885-329-6  
 ; Sequence 6, Application US/09885329  
 ; Patent No. US20020115170A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yaver, Debbie  
 ; APPLICANT: Mcardle, Barbara  
 ; TITLE OF INVENTION: Polypeptides Having Peroxidase Activity  
 ; FILE REFERENCE: 10048.200-US  
 ; CURRENT APPLICATION NUMBER: US/09/885,329  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: 09/596,824  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 362  
 ; TYPE: PRT  
 ; ORGANISM: Bjerkandera adusta  
 ; US-09-885-329-6

Query Match 8.3%; Score 132.5; DB 10; Length 362;  
 Best Local Similarity 24.8%; Pred. No. 8.4e-05;  
 Matches 84; Conservative 33; Mismatches 121; Indels 101; Gaps 16;

QY 14 SLTSCVA--HAQISPTFYASCPN-LOSTIVRAAMTQAVASEQRMASL----- 58  
 Db 8 SLTSAVAHIAAGALTRRVACPDGVNTATNAACPLIAYRDDMOXNLTDGACNAEVEH 67  
 QY 59 -LRLFHDCFVQ-----GCDGSIL---LDMGKETAQPNLNSVGEVIDT 100  
 Db 68 SLRLTFHDAIGSPALAAAGSFAGGADGSLTFPSDVEAFANALD-----EMIEL 120  
 QY 101 IKRIVER-ACPGVSCADILAAADG-TNLGGPTWSVPLGRDRSTASLANSP 157  
 Db 121 QKPYIKRYNTPG---DVQFAGAVGSLNCPGAPOLELIGR---TATAASPTGLIP 172  
 QY 158 PPTASCTLSLTFRQGLSPRDTALSGAHTIGARCTFRGRIVGDTINMSFALRQ 217  
 Db 173 APEDYDAITARRADVDSEVDELALLASHVAAS-----HIDT-----S 212  
 QY 218 TCRPSGGGNLADIDVOTPVREFPTAVFTNLISRRGLFH-----S 256  
 Db 213 TVPES-----PLD-STPGVFTQFVE-TSLNGTMYPTGTSIGNIGELSAIAGELRLLS 263  
 QY 257 DOELFNGSODALVROYASASLFNADFVAAMTRMGNG 295  
 Db 264 DHELARDSRTACEMOSTVNNQSIQSAFRAAMRAVIG 302

## RESULT 6

US-10-095-975-2  
 ; Sequence 2, Application US/10095975  
 ; Publication No. US20030013182A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yaver, Debbie S.  
 ; APPLICANT: Berka, Randy M.  
 ; TITLE OF INVENTION: Methods For Isolating Genes From  
 ; TITLE OF INVENTION: Microorganisms  
 ; FILE REFERENCE: 10160.200-US  
 ; CURRENT APPLICATION NUMBER: US/10/095,975  
 ; PRIOR FILING DATE: 2002-03-12  
 ; PRIOR APPLICATION NUMBER: 06/275,283  
 ; PRIOR FILING DATE: 2001-03-12

; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 376  
 ; TYPE: PRT  
 ; ORGANISM: Ceriportalopsis subvermispora  
 ; US-10-095-975-2

Query Match 7.9%; Score 126.5; DB 9; Length 376;  
 Best Local Similarity 23.4%; Pred. No. 0.00034;  
 Matches 60; Conservative 31; Mismatches 108; Indels 57; Gaps 11;

QY 2 ASPTLMQCL--VAVSLSCVAH---AQLSPFYASSC-PNLSIVRAAMTQAVASEQRM 54  
 Db 18 APTAV-CADGTRVNAACAPITPLAQLQELTFMGDCGEDAHEVIRLTFHDAVAISSM 76  
 QY 55 GASLRLRFHDCFVOGCDGSITLDAGGKTAGPNLNSVGEVIDTIRKNEAACPGVYS 114  
 Db 77 GPSA-----GGGADGSML---FPVPEPNSANNPID--DSVNNLIPFLSKHAVS 121  
 QY 115 CADILALARDG-TNLGGPTWSVPLGRDRSTASLANSPPTASLCTLSLTF-GR 172  
 Db 122 AGDLVQFAGAVALTNCPCGAPRLLEPLAGRNHTIA---IDGLIPEQDDVTKILAREDA 178  
 QY 173 QGLSPRDMTALSGAHTIGARCTFRGRIVGDTINASFALROOTCPRSGGDNLAIPID 232  
 Db 179 GGFSPFEVYSILASHTVARA-----DKVDGTIDNAP 209  
 QY 233 VOTPVREFPTAVFTNL 248  
 Db 210 DSTPFTFDTQVLEVL 225

## RESULT 7

US-09-748-264A-2  
 ; Sequence 2, Application US/09748264A  
 ; Patent No. US20020083492A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YOSHIKAWA, YOSHIO  
 ; APPLICANT: SECRETARY OF AGENCY OF INDUSTRIAL SCIENCE AND TECHNOLOGY  
 ; TITLE OF INVENTION: A BASIDIOMYCETE PEROXIDASE GENE-TRANSFERRED PLANT AND A METHOD  
 ; FILE REFERENCE: 04853.0050  
 ; CURRENT APPLICATION NUMBER: US/09/748,264A  
 ; PRIOR FILING DATE: 2000-12-27  
 ; PRIOR APPLICATION NUMBER: JP 2000-223653  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 365  
 ; TYPE: PRT  
 ; ORGANISM: Coriolus versicolor  
 ; US-09-748-264A-2

Query Match 7.8%; Score 125.5; DB 10; Length 365;  
 Best Local Similarity 25.1%; Pred. No. 0.00041;  
 Matches 74; Conservative 34; Mismatches 112; Indels 75; Gaps 15;

QY 1 MASPTLMQCLVAVSLSCVAHQAQISPTFYAS--SCPNI-LOSTIVRAAMTQAVASEQRM 57  
 Db 1 MAFKTL-----ASLISVLTIVQVAGALTRRVACPDGVNTATNAACCOLFAVHDDIQON 54  
 QY 58 L-----LRLFHDCFVOGCDGSITLDAGGKTAGPNLNSVGEVIDT----- 100  
 Db 55 LFDGECGEVHESIRLTFHAI--GISPSTA--SRQFGGGGADGSIALFEDIEITFNHA 110  
 QY 101 ---IKRNEAACPGV---VSCADILALARDG-TNLGGPTWSVPLGRDRSTASLANSP 152  
 Db 111 NLGVDEIDEGRPFIARHNTLADFIQFAGAIQVSNCPGAPQLDVLFIGRPDAPAPDL- 169  
 QY 153 NSNPPPTASLCTLSLTFG-RQGLSPRDMTALSGAHTIGARCTFRGRIVGDTINASF 211



Wed May 7 14:10:02 2003

us-10-047-825-4.rabp

Page 5

```
Db 115 GLEDEVAIAQK-----PFIAKHGVTGDFIAFAGVGVSCNPGAPOMQFELRPRPAQAA 168
OY 149 ASIANANPPPPASISGLTIS-LFGRGLSPRDMTALSGATTCQACTFFRRIRYGDPI 207
Db 169 ----PDQVPEPHTIDVLRKMDAGFDEIETWILSAHSIAAA-----NDV 213
OY 208 MASFAALRQQTCPRSSGSDGNALPDVQTVRDTAFYTNLISRRGIF 254
Db 214 DPTISGL-----PFQDGFDSQFVEVETQLKRTAF 244

RESULT 11
US-09-981-900B-14
Sequence 14, Application US/09981900B
GENERAL INFORMATION:
APPLICANT: Masomeh B
APPLICANT: Masomeh B
APPLICANT: Masomeh B
TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULOSE WHICH DEGRAD
FILE REFERENCE: MSU 4.1-539
CURRENT APPLICATION NUMBER: US/09/981,900B
PRIOR FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 60/242,408
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 14
LENGTH: 371
TYPE: PRT
ORGANISM: Phanerochaete chrysosporium
US-09-981-900B-14

Query Match
Best Local Similarity 22.6%; Score 89.5; DB 10; Length 371;
Matches 57; Conservative 32; Mismatches 88; Indels 75; Gaps 14;

OY 59 LRLFFHDCP-----VOCDCGSLILDAGSEKTAGENLNSVGFVDTJTKR-- 103
Db 69 IRLVFHDAIAISPAWEPQASSVRGADGSIIMIDEIETNPHNTI---GIDEIVRLQKPFV 124
OY 104 NVEACPGVSCADITLALARDG-TNLTGPTWVPLGRDSTTASASLANSNPPPTAS 162
Db 125 OKHGVPFG-----DEIAFAGVVALSNCPGAPOMQFTRAPATOPA---PDGLVPEPFS 176
OY 163 LGLTIS-LFGRGLSPRDMTALSGATTCQACTFFRRIRYGDPIINASFAALRQOTCPR 221
Db 177 VDQIIRVFDAGEFDELELVWMLSAHSVAAA-----NDIDPNIQGL----- 217
OY 222 SGGGMIAPIDVQTVRDTAFYTNL-LSRRLPHSDQELFNGSGSDALVROYASASL- 279
Db 218 -----PFD-STPGIRDSQFVEVETQLKRTAF-----PTGSSNN---QGEVSPLP 256
OY 280 ----FNADVAA 287
Db 257 GEMRLQSDFLIA 268

RESULT 12
US-10-101-464A-802
Sequence 802, Application US/10101464A
GENERAL INFORMATION:
APPLICANT: Stradala, Timothy
APPLICANT: Neuenhulzen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions isolated from plant cells
FILE REFERENCE: 11000,1020C2
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: US/10/101,464A
PRIOR FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
```

```
Db 115 GLEDEVAIAQK-----PFIAKHGVTGDFIAFAGVGVSCNPGAPOMQFELRPRPAQAA 168
OY 149 ASIANANPPPPASISGLTIS-LFGRGLSPRDMTALSGATTCQACTFFRRIRYGDPI 207
Db 169 ----PDQVPEPHTIDVLRKMDAGFDEIETWILSAHSIAAA-----NDV 213
OY 208 MASFAALRQQTCPRSSGSDGNALPDVQTVRDTAFYTNLISRRGIF 254
Db 214 DPTISGL-----PFQDGFDSQFVEVETQLKRTAF 244

RESULT 11
US-09-981-900B-14
Sequence 14, Application US/09981900B
GENERAL INFORMATION:
APPLICANT: Masomeh B
APPLICANT: Masomeh B
APPLICANT: Masomeh B
TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULOSE WHICH DEGRAD
FILE REFERENCE: MSU 4.1-539
CURRENT APPLICATION NUMBER: US/09/981,900B
PRIOR FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 60/242,408
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 14
LENGTH: 371
TYPE: PRT
ORGANISM: Phanerochaete chrysosporium
US-09-981-900B-14

Query Match
Best Local Similarity 22.6%; Score 89.5; DB 10; Length 371;
Matches 57; Conservative 32; Mismatches 88; Indels 75; Gaps 14;

OY 59 LRLFFHDCP-----VOCDCGSLILDAGSEKTAGENLNSVGFVDTJTKR-- 103
Db 69 IRLVFHDAIAISPAWEPQASSVRGADGSIIMIDEIETNPHNTI---GIDEIVRLQKPFV 124
OY 104 NVEACPGVSCADITLALARDG-TNLTGPTWVPLGRDSTTASASLANSNPPPTAS 162
Db 125 OKHGVPFG-----DEIAFAGVVALSNCPGAPOMQFTRAPATOPA---PDGLVPEPFS 176
OY 163 LGLTIS-LFGRGLSPRDMTALSGATTCQACTFFRRIRYGDPIINASFAALRQOTCPR 221
Db 177 VDQIIRVFDAGEFDELELVWMLSAHSVAAA-----NDIDPNIQGL----- 217
OY 222 SGGGMIAPIDVQTVRDTAFYTNL-LSRRLPHSDQELFNGSGSDALVROYASASL- 279
Db 218 -----PFD-STPGIRDSQFVEVETQLKRTAF-----PTGSSNN---QGEVSPLP 256
OY 280 ----FNADVAA 287
Db 257 GEMRLQSDFLIA 268

RESULT 12
US-10-101-464A-802
Sequence 802, Application US/10101464A
GENERAL INFORMATION:
APPLICANT: Stradala, Timothy
APPLICANT: Neuenhulzen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions isolated from plant cells
FILE REFERENCE: 11000,1020C2
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: US/10/101,464A
PRIOR FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302

PRIORITY FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12/162,866
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724
PRIOR FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 802
LENGTH: 345
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-10-101-464A-802

Query Match
Best Local Similarity 22.8%; Score 89; DB 9; Length 345;
Matches 65; Conservative 33; Mismatches 107; Indels 80; Gaps 12;

OY 9 CIVAASLSCVAAQOLPTFYVASSCPNOSIVRAAMTQAVASEQPMGASLRLP-PRDCF 67
Db 62 CSAAMRGVRCSSSTHSSRVATLSTPSLN-LRGPIDSLSDQ-----LRLDHDNR 113
OY 68 VOC-----CDGSLILDAGSEKTAG--PNTSVGFVDTJTKRNVAAAGVSCAD 117
Db 114 LNGVYTPLANCTNMLKLYLAGNDFSGEIPPEISLRRLRDISNNIRVYIGEVRGIS 173
OY 118 ILAALARDGTNLLG--PTWSVPL-----GRR----- 142
Db 174 RL-LTLRQSLSGEVPDLVSLKLNKELNLTNNELTGHLPDGLKRFVRSFVNGECL 232
OY 143 -DSTTASLANSNP-----PPTASLGLTISLF--KQGLSPDMTA 182
Db 233 GSSPLACSLSGTSPASSSLTVPSNPSMPETPAIG--GQKHQAGSTGATIVA 287
OY 183 ISGATTCQACTFFRRIRYGDPIINASFAALRQOTCPRSSGSDGN 227
Db 288 IALNCVALLVTSFVAVCCSRHNSSLKANGESTGKRSSSH 332

RESULT 13
US-09-782-906-5
Sequence 5, Application US/09782906
Patent No. US20010051369A1
GENERAL INFORMATION:
APPLICANT: Delagrave, Simon
APPLICANT: Rittenhouse Pruss, Jennifer L.
APPLICANT: Murphy, Dennis J.
APPLICANT: Maffia III, Anthony M.
APPLICANT: Bylina, Edward J.
TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Met
FILE REFERENCE: HBR-0040
CURRENT APPLICATION NUMBER: US/09/782,906
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/185,001
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 5
LENGTH: 639
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: No. US20010051369A1 Sequence
US-09-782-906-5

Query Match
Best Local Similarity 20.5%; Score 84.5; DB 10; Length 639;
Matches 63; Conservative 24; Mismatches 95; Indels 125; Gaps 13;

OY 64 HDCEVQCG-----DGSILLDAG--GKTRA-----GPMLSVGRFVEVDPI----- 101
```

Wed May 7 14:10:02 2003

us-10-047-825-4.rapb

Page 6

Db 224 HOMFCGISMONGOIYVGTGNDAXERTSLYSSDSMTIPCPDQVARGVOSATFMSDGRV 283  
QY 102 -----KRVNAACP-----GVSCADILAAAP-----G 126  
Db 284 FTIGSMGCVFEKNGEVSPSSKWTSLPNAKYNPHLITDKGCLRSBNHAWLEGMKKG 343  
QY 127 TNLGGPT-----NSVPLGRDSTTASASLANSNPPT----- 160  
Db 344 SVQAFPSAMNWTYSOSDVAACKROSNKVAFPMAGMAGVMDAKKILTFGGSP 403  
QY 161 -----ASLGTLSLEGRGCLSPRDMTALSGAHTIGARCTTFRRIY---GOTDINASF 211  
Db 404 DVODSDATTNAHITTEDEPTSPNYVFASNGILF-----ARHTSVLPLDSTFT--- 454  
QY 212 AALROQTPRSGGDNAPIDVOTPV-----RFTYAFYMLSRRLFSH-----D 257  
Db 455 -----TGGRGRIPEDESTPVETPEIYVEODTFYKONPSIVRAVHSISLLPD 504  
QY 258 QELFNGG 264  
Db 505 GRVFNNG 511

RESULT 14  
US-10-101-464A-890  
Sequence 890, Application US/10101464A  
Publication No. US20030046728A1  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
APPLICANT: Higgins, Colleen M.  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TIME OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
FILE REFERENCE: 11000.102062  
CURRENT APPLICATION NUMBER: US/10/101.464A  
CURRENT FILING DATE: 2002-03-18  
PRIOR APPLICATION NUMBER: 09/704.302  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 09/228.986  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/162.866  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: PCT/US00/00724  
PRIOR FILING DATE: 2000-01-11  
NUMBER OF SEQ ID NOS: 989  
SOFTWARE: FastSeq for Windows Version 4.0.  
SEQ ID NO 890  
LENGTH: 804  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-10-101-464A-890

Query Match 5.3%; Score 84.5; DB 9; Length 804;  
Best Local Similarity 22.6%; Pred. No. 12;  
Matches 60; Conservative 37; Mismatches 122; Indels 47; Gaps 11;

QY 26 PTFYASSCPNQSIVRAA--MTQAVASEQRMGASILNLFRRDFVVG-----CDGSI 75  
Db 234 PYSLA-NCTSLDLVYSSNNLSGPISSESGVSKFLFDNDSISGSTPISVNCYSLF 292  
QY 76 LLDAGERTAGPNNLSVYG---FEVIDTIRKNEAACP-GVSCADILAAARDGTNLL 130  
Db 293 ALKGSQNKLSGP-IPSVGMGLKXSIYTDLSNNSLGRLPTTLVNCQYELYL--STNNL 349  
QY 131 GGFYWSVPLGRDSTTASASLANS---NPPPTASLGTLSL----- 169  
Db 350 RS-SISIPSTIDTDLVYLDLSNNQLIGSLPNIIGSLQSLVLSNNNLGSLPHNIQ 408  
QY 170 ---FGRGSLSPDMTA---LSGAHTIGQARCTFRGRIRYGDIDINAFALROQCPRG 223  
Db 409 LSDLQRLDLSNNNLGSLPLSISHIEDLOEDLYNNNIFGTINPTISLTSIQIIDLSKNN 468

QY 224 GDGNLAPIDVOTVPRFDTAYFTNLIS 249  
Db 469 LDGSIPELNSCEPELSFFSVHTNKLIS 494

RESULT 15  
US-09-976-059-14  
Sequence 14, Application US/09976059  
Patent No. US20020164747A1  
GENERAL INFORMATION:  
APPLICANT: Farnec, Chris  
APPLICANT: Zazopoulos, Emmanuel  
APPLICANT: Stafla, Alfredo  
TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin  
FILE REFERENCE: 3019-PCT  
CURRENT APPLICATION NUMBER: US/09/976.059  
CURRENT FILING DATE: 2001-10-15  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 14  
LENGTH: 4999  
TYPE: PRT  
ORGANISM: Actinoplanes sp.  
US-09-976-059-14

Query Match 5.2%; Score 83.5; DB 9; Length 4999;  
Best Local Similarity 27.9%; Pred. No. 1.9e+02;  
Matches 58; Conservative 21; Mismatches 94; Indels 35; Gaps 9;

QY 103 RNVAACPGVSCADILAAARDGTNLSGPTWSVPLGRDSTTASASLAN---SNPPP 159  
Db 1740 RVAGFAPGDMVDKSMAL-----LAGAALFFVEPERRGADLAFLAVGVTHATLP 1794  
QY 160 TASGTLISLEGRGCLSPRDMTALSGAHTI--GARCTTFRGRI---YGDIDINAF 213  
Db 1795 PAVVGT-----PDGVLPSYLDVGDVLPGLAKRMILROGRVLFNSYGTETTYNNAF 1848  
QY 214 LRQOTCRSGGDCNLAVIDVOTPVPRFDTAYFTNLISRLGSLHSDQLFNGSQDALVRY 273  
Db 1849 WRAE---AGDMSVAPLG--TPVNNLRATVLDGMLRPVPGADGDELVSQA--GLARGI 1900  
QY 274 SASASLFNADVAA-----MIRGNV 294  
Db 1901 LNRAGLTAERFVACPEPEGERMYRTGDV 1928

Search completed: April 26, 2003, 12:43:24  
Job time : 46 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 3, 2003, 13:16:38 ; Search time 2174 Seconds

(without alignments)  
4190.054 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 313

Sequence: 1 MASPTLMQCLVAVSLISCAV.....VGVLTGTAGQIRNCRVNVN 313

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2054640 seqs, 14551402878 residues

Word size: 16

Total number of hits satisfying chosen parameters: 55

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODL-frame+ p2n.model -DEV-xlh  
-O/cgn2\_1/USPTO.spool/US10047825/runat\_26042003\_112042\_7644/app\_query.fasta.1.455  
-DB-GenEmbl -QFMT-fastap -SUFFIX-rge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=0.190 -TRANS-human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=qualify -THR MIN=16 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro  
-NORR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10047825\_GCGN.1.1.1687\_6runat.26042003\_112042\_7644 -NCPU=3  
-NO\_XLPRY -NO\_WMAP -LARGEQUERY -NES\_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rnd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	8.3	1335	8	CCU12314
2	21	6.7	160562	2	AP005613
3	20	6.4	835	8	AF149279
4	20	6.4	870	8	ATR6961
5	20	6.4	960	6	AX412410
6	20	6.4	960	6	AX412799
7	20	6.4	965	8	ATP24A
8	20	6.4	987	6	AX412383
9	20	6.4	987	6	AX412578
10	20	6.4	1271	8	AY086282
11	20	6.4	1287	8	AF428274
12	20	6.4	1330	8	ARCPCN2
13	20	6.4	1389	8	AF485268
14	20	6.4	1400	8	ATP02
15	20	6.4	48622	8	AB008266
16	20	6.4	68697	2	AB012243
17	20	6.4	158456	2	AP004731
18	18	5.8	1060	8	ALPFXD
19	18	5.8	1658	8	TAFOX2
20	18	5.8	3240	8	TAPEROXIG
21	18	5.8	201642	2	AP005529
22	17	5.4	254	8	AF067188
23	17	5.4	993	8	SORRXR7
24	17	5.4	1083	8	ATRAP20A
25	17	5.4	1139	8	GM051194
26	17	5.4	1160	8	LECEVIL16G
27	17	5.4	1167	6	AR032064
28	17	5.4	1167	6	AR060347
29	17	5.4	1172	8	AB027753
30	17	5.4	1190	8	AY057607
31	17	5.4	1222	8	AF442386
32	17	5.4	1369	8	AF155124
33	17	5.4	1369	8	SSNCAPE
34	17	5.4	42499	8	AC004625
35	17	5.4	65465	8	AB007650
36	17	5.4	84987	2	AC123571
37	17	5.4	105692	2	AP003953
38	17	5.4	107424	2	AP005290
39	17	5.4	135792	2	AP003544
40	17	5.4	144743	2	AP004674
41	17	5.1	957	8	AB024439
42	16	5.1	1172	8	GM051193
43	16	5.1	1176	8	AF049881
44	16				
45	16				

RESULT 1

#### ALIGNMENTS

CCU12314 1335 bp mRNA linear PLN 02-AUG-1994  
 CCU12314 Cenchrus ciliaris clone PX7 peroxidase mRNA, complete cds.  
 DEFINITION U12314  
 ACCESSION U12314.1 GI:520567  
 VERSION  
 KEYWORDS  
 SOURCE Cenchrus ciliaris.  
 ORGANISM Cenchrus ciliaris.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Paniceae; Cenchrus.  
 1 (bases 1 to 1335)  
 REFERENCE  
 AUTHORS Ross, A.H.  
 TITLE Investigation of peroxidase genes and genetic transformation in  
 JOURNAL buffel grass  
 REFERENCE 2 (bases 1 to 1335)  
 AUTHORS Ross, A.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-JUL-1994) Annette H. Ross, University of Queensland,  
 Botany, St. Lucia, Brisbane, Queensland, 4072, Australia  
 FEATURES  
 source  
 1. 1335  
 /organism="Cenchrus ciliaris"  
 /cultivar="Biloele"  
 /db\_xref="taxon:35872"  
 /clone="PX7"  
 /tissue\_type="wounded and unwounded leaves and stems"  
 /clone\_1lb="lambda gt-10"  
 1. 57  
 58. 981  
 /EC\_number="1.11.1.7"  
 /note="expressed in leaves and is wound-induced"  
 /codon\_start=1  
 /product="peroxidase"  
 /protein\_id="AAA20472.1"  
 /db\_xref="GI:520568"  
 /translation="MAAPFLMQCLALSLSTFHAQLSTFTYASCPNLCQTVRAAM  
 TOVSSSEPRKASLIRLFHDFVQGCDASTLSRSGESAGPNANSVRGEVIDTK  
 KNEVACPGVVSCTIVPLARPGPNLGGFTWVPLGRDSTTAMSLANONLPPPT  
 SLGTLISLFGRLSARDMLALSGAHQARCTFRGRIYGTNTIDASFALQOOTCR  
 SGSDNLAPIDACTPARDNAYTNILVSRGLFHSDDLEFNGSDALVHYSSPSQ  
 FNSDEVAMIKMIGNAGGVRNCRVYNS"  
 58. 126  
 /note="putative"  
 mat\_peptide 127..978  
 /product="peroxidase"  
 3'UTR 982..1335  
 polyA\_site 1335  
 /note="39 A residues"  
 BASE COUNT 264 a 440 c 374 g 257 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9.69e-18 Length: 1335  
 Score: 26.00 Matches: 26  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 8.31% Indels: 0  
 DB: Gaps: 0  
 US-10-047-825-4 (1-313) x CCU12314 (1-1335)  
 QY 249 SerAARGAGGyLeuPheHisSerAspIngluLeuPheAsnglyGlySerGlnAspAla 268  
 |||||  
 DB 793 TCGCGGCGGGGCTTTTCACGACGAGAGCTTCAACGGCGGTCGACAGAGCGG 852  
 |||||  
 QY 269 LeuValArgGlnTySer 274  
 |||||  
 DB 853 CTGGTAGGACACTACAGC 870  
 |||||  
 RESULT 2  
 AP005613/c

LOCUS AP005613 160562 bp DNA linear HTG 02-AUG-2002  
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone  
 OSJNBa0082C09, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.  
 ACCESSION AP005613  
 VERSION AP005613.1 GI:22091085  
 KEYWORDS HTG; HTGS; PHASE2.  
 SOURCE Oryza sativa (japonica cultivar-group) (cultivar: Nipponbare) DNA,  
 clone:OSJNBa0082C09.  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1  
 REFERENCE  
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.  
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC  
 clone:OSJNBa0082C09  
 JOURNAL Published Only in Database (2002)  
 REFERENCE 2 (bases 1 to 160562)  
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-AUG-2002) Takuji Sasaki, National Institute of  
 Agricultural Sciences, Rice Genome Research Program, Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/,  
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)  
 NOTE: It currently consists of 1 contigs. Gaps between the contigs  
 are represented as runs of N. The order of the pieces is believed  
 to be correct as given, however the sizes of the gaps between them  
 are based on estimates that have provided by the submitter. This  
 sequence will be replaced by the finished sequence as soon as it is  
 available and the accession number will be preserved.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 FEATURES  
 source  
 1. 160562  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="2"  
 /clone="OSJNBa0082C09"  
 BASE COUNT 45817 a 33079 c 35854 g 45760 t 52 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.61e-10 Length: 160562  
 Score: 21.00 Matches: 21  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.71% Indels: 0  
 DB: Gaps: 0  
 US-10-047-825-4 (1-313) x AP005613 (1-160562)  
 QY 104 AsnValGlnAlaAlaGcSPGGLyValSerCysAlaAspIleLeuAlaLeuAla 123  
 |||||  
 DB 77736 AACGTCGAGCGCGCGCGCGCGCTCTCTCGCGCGACATCTCGCGCGCGC 77677  
 |||||  
 QY 124 Arg 124  
 |||  
 DB 77676 CGC 77674  
 |||  
 RESULT 3  
 AF149279  
 LOCUS AF149279 835 bp mRNA linear PLN 14-DEC-1999  
 DEFINITION Phaseolus vulgaris peroxidase 4 precursor (FBP4).mRNA, partial cds.  
 ACCESSION AF149279  
 VERSION AF149279.2 GI:6573300  
 KEYWORDS  
 SOURCE Phaseolus vulgaris.  
 ORGANISM Phaseolus vulgaris.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.

REFERENCE 1 (bases 1 to 835)  
AUTHORS Blee,K.A., Jupe,S.C., Richard,G. and Bolwell,G.P.  
TITLE Molecular cloning and differential expression of six members of the peroxidase family in French bean  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 694)  
AUTHORS Blee,K.A. and Bolwell,G.P.  
TITLE Direct Submission  
JOURNAL Submitted (10-MAY-1999) Biochemistry, Royal Holloway, University of London, Egham, Surrey TW20 OEX, United Kingdom  
REFERENCE 3 (bases 1 to 835)  
AUTHORS Blee,K.A. and Bolwell,G.P.  
TITLE Direct Submission  
JOURNAL Submitted (14-DEC-1999) Biochemistry, Royal Holloway, University of London, Egham, Surrey TW20 OEX, United Kingdom  
REMARK Sequence update by submitter  
COMMENT On Dec 14, 1999 this sequence version replaced gi:5002345.  
FEATURES  
source  
1. 835  
/organism="Phaseolus vulgaris"  
/cultivar="Moncalm"  
/db\_xref="taxon:3885"  
/clone="PB750-04.5"  
/c1.>835  
/gene="FBP4"  
/note="secretory peroxidase"  
/codon\_start=2  
/product="peroxidase 4 precursor"  
/protein\_id="A037429.2"  
/db\_xref="GI:6573301"  
/translation="FYSSSCPRAESIVKSTVSHVKSSTLAAGLLRMHFDCEYQGC  
DGSVLLSGANTERTAFANGLRGFEVVDKATOLENACPGVSCADILALARDSVL  
SGSLSYQVPTGRDGRISQADSVSNLPAPDSVDYOKKFTAKGLNTQDLVLLGHT  
IGTTACOFENRLNYNFTANGPDSIDSPFLPTQSCPNQDSTVALDTSOKLFD  
LSYNNIRKRGRLIQSDQALMSDSDSKYVQRTGLIRGLIGLKFVEEENGMVMKKNG  
IG"  
BASE COUNT 197 a 193 c 199 g 246 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.78e-11 Length: 835  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.39% Indels: 0  
Gaps: 0  
DB: 8  
US-10-047-825-4 (1-313) x AF149279 (1-835)  
QY 106 GUAAlaAcYProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAsp 125  
|||||  
Db 236 GAGGCCGCGATCCCGGTCGTCTGTCTGCTGATATCTTGCCTTGGCTGCTGAT 295  
|||||  
RESULT 4  
ATH6961 870 bp DNA linear PLN 16-DEC-1998  
LOCUS Arabidopsis thaliana p15a gene, partial.  
ACCESSION AJ006961  
VERSION AJ006961.1 GI:3242069  
KEYWORDS p15a gene; peroxidase.  
SOURCE  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 870)  
AUTHORS Justesen,A.F.  
TITLE Direct Submission

JOURNAL Submitted (15-JUN-1998) Justesen A.F., Department of Protein Chemistry, Institute of Molecular Biology, University of Copenhagen, Øster Farimagsgade 2A, 1353 Copenhagen K, DENMARK  
REFERENCE 2 (bases 1 to 870)  
AUTHORS Justesen,A.F., Jespersen,H.M. and Welinder,K.G.  
TITLE Analysis of two incompletely spliced Arabidopsis cDNAs encoding novel types of peroxidase  
JOURNAL Biochim. Biophys. Acta 1443 (1-2), 149-154 (1998)  
MEDLINE 99057591  
PUBMED 9838086  
FEATURES  
source  
Location/Qualifiers  
1. 870  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
2. 870  
/gene="p15a"  
join(<2..277,387..549,850..>870)  
/gene="p15a"  
join(<2..277,387..549,850..>870)  
/gene="p15a"  
/EC\_number="1.11.1.7"  
/codon\_start=1  
/product="peroxidase"  
/protein\_id="CA07353.1"  
/db\_xref="GI:4375832"  
/translation="TIVNAYAGSSDPRIAPGILRMHFDCEYQGCDSILISGAN  
TERTRAGPRLNAGFEVDNKTOLKACPGVSCADILALARPDTVLITGTGQVPT  
GRDRKVSILASNNANLPGRDVSVAQOKFSALNTRDLVLYVGHHTIG"  
2. 277  
/gene="p15a"  
/number=1  
278..386  
/gene="p15a"  
/number=1  
387..549  
/gene="p15a"  
/number=2  
550..849  
/gene="p15a"  
/number=2  
850..870  
/gene="p15a"  
/number=3  
BASE COUNT 261 a 190 c 150 g 269 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.85e-11 Length: 870  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.39% Indels: 0  
Gaps: 0  
DB: 8  
US-10-047-825-4 (1-313) x ATH6961 (1-870)  
QY 106 GUAAlaAcYProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAsp 125  
|||||  
Db 206 GAGGCCGCGGTCTGAGAGTGTCTGTCTGCTGATATTTCATTAACCCCTGTCGAC 265  
|||||  
RESULT 5  
AX412410 960 bp DNA linear PAT 14-JUN-2002  
LOCUS Sequence 174 from Patent WO0222675.  
ACCESSION AX412410  
VERSION AX412410.1 GI:21444868  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

	REFERENCE	1	Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
	AUTHORS		Plant genes, the expression of which are altered by pathogen infection
	TITLE		
	JOURNAL		Patent: WO 0222675-A 174 21-MAR-2002;
			Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl, Jeffrey L. (US) ; Eulgem, Thomas (US)
	FEATURES		Location/Qualifiers
	source	1..960	
		/organism="Arabidopsis thaliana"	
		/db_xref="taxon:3702"	
	BASE COUNT	246 a 225 c 229 g 260 t	
	ORIGIN		
	Alignment Scores:		
	Pred. No.:	2.02e-11	Length: 960
	Score:	20.00	Matches: 20
	Percent Similarity:	100.00%	Conservative: 0
	Best Local Similarity:	100.00%	Mismatches: 0
	Query Match:	6.39%	Indels: 0
	DB:	6	Gaps: 0
	US-10-047-825-4 (1-313) x AX412799 (1-960)		
OY	106 GIUAlAalAcysPProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgasp	125	
Db	319 GAGCGACATGTCTCGTGGTGTTGTCCTTGCTCATATCTTGCGCTTAGCGGCTCGTAT	378	
RESULT 6	AX412799	960 bp	DNA linear PAT 14-JUN-2002
LOCUS	AX412799		
DEFINITION	Sequence 563 from Patent W00222675.		
ACCESSION	AX412799		
VERSION	AX412799.1 GI:21445257		
KEYWORDS			
SOURCE			
ORGANISM	thale cress.		
	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
REFERENCE	1		
AUTHORS	Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.		
TITLE	Plant genes, the expression of which are altered by pathogen infection		
JOURNAL	Patent: WO 0222675-A-563 21-MAR-2002;		
	Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl, Jeffrey L. (US) ; Eulgem, Thomas (US)		
FEATURES	Location/Qualifiers		
source	1..960		
	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
BASE COUNT	246 a 225 c 229 g 260 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.02e-11	Length: 960	
Score:	20.00	Matches: 20	
Percent Similarity:	100.00%	Conservative: 0	
Best Local Similarity:	100.00%	Mismatches: 0	
Query Match:	6.39%	Indels: 0	
DB:	6	Gaps: 0	
US-10-047-825-4 (1-313) x AX412799 (1-960)			
OY	106 GIUAlAalAcysPProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgasp	125	
Db	319 GAGCGACATGTCTCGTGGTGTTGTCCTTGCTCATATCTTGCGCTTAGCGGCTCGTAT	378	
RESULT 7	ATP24A	965 bp	mRNA linear PLN 14-MAR-1997
ATP24A			
LOCUS			

DEFINITION	A.thaliana mRNA for peroxidase AtP24a, partial.
ACCESSION	Y11788
VERSION	Y11788.1 GI:1890312
KEYWORDS	peroxidase.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
AUTHORS	1 (bases 1 to 965) Wellinder,K.G., Jespersen,H.M., Kjaergaard,I.V.M., Justesen,A.F., Oestergaard,L., Abelskov,A.K., Jensen,R.B., Hansen,L.N. and Raasmussen,S.K.
TITLE	From expressed sequence tags to structure, function, evolution and expression of 28 ER-targeted Arabidopsis peroxidases
JOURNAL	Unpublished 2 (bases 1 to 965)
REFERENCE	Justesen,A.F. Direct Submission Submitted (11-MAR-1997) A.F. Justesen, University of Copenhagen, Dept. of Protein Chemistry, Oester Farimagsgade 2a, DK-1353 Copenhagen K, DENMARK
AUTHORS	Location/Qualifiers
JOURNAL	1..965
REFERENCE	/organism="Arabidopsis thaliana"
AUTHORS	/db_xref="taxon:3702"
TITLE	/clone="VBVMG04"
JOURNAL	/cell_line="Columbia"
REFERENCE	/clone_id="Versailles-VB"
AUTHORS	/dev_stage="retiolated seedlings"
TITLE	<1..774
JOURNAL	/EC_number="1.11.1.7"
REFERENCE	/codon_start=1
AUTHORS	/product="peroxidase AtP24a"
TITLE	/protein_id="CAAT2484.1"
JOURNAL	/db_xref="GI:1890313"
REFERENCE	/db_xref="SPTREMBL:P93724"
AUTHORS	/translation="HNHDCFEVQCDGSVLSPNSERTAGANNVLHGFEVIDDAKROIT EAACPGVASCADIIALAAARDSVSLTNGQSWVPYGRDGRVSLASNNNLPSPDSLSL IQOKRFAFRLNTDLVTLVGGGHITIGTAACGFITNRIFNSGNTPADPTMDOTVEPOL QRCPQMGDSGARVDLTGTSGNFETSTFIILSNRRGLIOSEHLMTSPATRSIVQEEL MAPRGNNVCARFARMKMSNIGTVTGNGEIRRCVSAVN"
TITLE	BASE COGNT 275 a 199 c 218 g 273 t
JOURNAL	ORIGIN
REFERENCE	Alignment Scores:
AUTHORS	Pred. NO.: 2.03e-11 Length: 965
TITLE	Score: 20.00 Matches: 20
JOURNAL	Percent Similarity: 100.00% Conservative: 0
REFERENCE	Best local Similarity: 100.00% Mismatches: 0
AUTHORS	Query Match: 6.39% Indels: 0
TITLE	DB: 8 Gaps: 0
JOURNAL	STS-10-047-825-4 (1-313) x ATP24A (1-965)
REFERENCE	OY 106 GIUAAlaAcysProGlyValValsercysylaAspileuAlalaLeuAlaAlaargasp 125
AUTHORS	 
TITLE	Dn 133 GAGCGAGCATCTCCTGCTGTTCCTTTCTGTCTATATCTTGCCCTTAGCGGCTCGTAT 192
JOURNAL	RESULT 8
REFERENCE	AX412383
AUTHORS	LOCUS AX412383
TITLE	DEFINITION Sequence 147 from Patent WO0222675.
JOURNAL	AX412383
REFERENCE	AX412383.1 GI:21444841
AUTHORS	KEYWORDS
TITLE	SOURCE
JOURNAL	ORGANISM
REFERENCE	thale cress. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
AUTHORS	Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.

TITLE Plant genes, the expression of which are altered by pathogen infection

JOURNAL Patent: WO 0222675-A 147 21-MAR-2002;  
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl, Jeffrey L. (US) ; Eulgem, Thomas (US)  
Location/Qualifiers

FEATURES  
source 1..987  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"

BASE COUNT 245 a 256 c 238 g 248 t

ORIGIN

Alignment Scores:  
Pred. No.: 2.07e-11 Length: 987  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.39% Indels: 0  
DB: Gaps: 0

US-10-047-825-4 (1-313) x AX412383 (1-987)

OY 106 GluAlaAlaCysProGlyValValSerCysAlaAspIleuAlaLeuAlaAlaArgasp 125  
Db 349 GAAGCCGCGTCTGAGATTGCTGTGCTGATATTAGCATTTAGCCGCTCGTGAC 408

RESULT 9  
AX412578

LOCUS AX412578 987 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 342 from Patent W00222675.  
ACCESSION AX412578  
VERSION AX412578.1 GI:21445036  
KEYWORDS  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE  
AUTHORS 1  
TITLE Glazebrook, J., Wang, X., Dangl, J.L., Eulgem, T. and Zhu, T.  
JOURNAL Plant genes, the expression of which are altered by pathogen infection  
PATENT: WO 0222675-A 342 21-MAR-2002;  
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl, Jeffrey L. (US) ; Eulgem, Thomas (US)  
Location/Qualifiers

FEATURES  
source 1..987  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"

BASE COUNT 245 a 256 c 238 g 248 t

ORIGIN

Alignment Scores:  
Pred. No.: 2.07e-11 Length: 987  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.39% Indels: 0  
DB: Gaps: 0

US-10-047-825-4 (1-313) x AX412578 (1-987)

OY 106 GluAlaAlaCysProGlyValValSerCysAlaAspIleuAlaLeuAlaAlaArgasp 125  
Db 349 GAAGCCGCGTCTGAGATTGCTGTGCTGATATTAGCATTTAGCCGCTCGTGAC 408

RESULT 10  
AX412813

LOCUS AX412813 987 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 577 from Patent W00222675.  
ACCESSION AX412813

VERSION AX412813.1 GI:21445271  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana  
thale cress.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE  
AUTHORS 1  
TITLE Glazebrook, J., Wang, X., Dangl, J.L., Eulgem, T. and Zhu, T.  
JOURNAL Plant genes, the expression of which are altered by pathogen infection  
PATENT: WO 0222675-A 577 21-MAR-2002;  
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl, Jeffrey L. (US) ; Eulgem, Thomas (US)  
Location/Qualifiers

FEATURES  
source 1..987  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"

BASE COUNT 245 a 256 c 238 g 248 t

ORIGIN

Alignment Scores:  
Pred. No.: 2.07e-11 Length: 987  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.39% Indels: 0  
DB: Gaps: 0

US-10-047-825-4 (1-313) x AX412813 (1-987)

OY 106 GluAlaAlaCysProGlyValValSerCysAlaAspIleuAlaLeuAlaAlaArgasp 125  
Db 349 GAAGCCGCGTCTGAGATTGCTGTGCTGATATTAGCATTTAGCCGCTCGTGAC 408

RESULT 11  
AY086282

LOCUS AY086282 1271 bp mRNA linear PLN 26-JUN-2002  
DEFINITION Arabidopsis thaliana clone 23349 mRNA, complete sequence.  
ACCESSION AY086282  
VERSION AY086282.1 GI:21404992  
KEYWORDS  
SOURCE Full-length messenger RNA sequences greatly improve genome annotation.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE  
AUTHORS 1  
TITLE Haas, B.J., Volfovsky, N., Town, C.D., Troupkhan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.  
JOURNAL Full-length messenger RNA sequences greatly improve genome annotation  
Genome Biol. (2002) In press  
2 (bases 1 to 1271)  
Brover, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.  
Full-length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 1271)  
Brover, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.  
Direct Submission  
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA  
This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be



AUTHORS Buffard, D., Breda, C., van Huystee, R. B., Asemota, O., Pierre, M., Ha, D. B. and Esnault, R.  
 TITLE Molecular cloning of complementary DNAs encoding two cationic peroxidases from cultivated peanut cells  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (22), 8874-8878 (1990)  
 MEDLINE 91062381  
 PUBMED 2247460  
 COMMENT Draft entry and computer-readable sequence for (unpublished (1990)) kindly submitted by R. Esnault, 08-AUG-1990.

FEATURES  
 SOURCE Institut des Sciences Vegetales  
 C.N.R.S.  
 91958 Clif Sur Yvette, Cedex.  
 location/Qualifiers

gene  
 1..1330 /organism="Arachis hypogaea"  
 /db\_xref="taxon:3818"  
 1..1330 /gene="PNC2"  
 <1..1330 /gene="PNC2"  
 200..1192 /gene="PNC2"  
 /codon\_start=1  
 /product="cationic peroxidase"  
 /protein\_id="AA32676.1"  
 /db\_xref="GI:166475"  
 /translation="MEGVFNKKFLIVEFMALGICIGITVTHGQTRVGYRSCPPA  
 ESIYRSTVSHVNSDPTLAKILRMHFDGFCVGCDSITLISGPTERTAFANGLRG  
 YEIDDAKTOLEACPGVASCADILALAAKDSVLSGSLSMQVPRGRDVSQASDV  
 SNLPAPSDVQOKFAKGLNTQDITVLTGHTITGTECECFENRILFENGTAAAD  
 PAIDPSVSLQALCPQNGKANGANRVALDGTGSRPFSVSNLRNRGVLQDQALMN  
 DPTKSFVQVRLGIRGLGTFNIEFGKSNMKNISGVTGIDGIRIKCARF"

BASE COUNT 354 a 296 c 256 g 424 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.69e-11 Length: 1330  
 Score: 20.00 Matches: 20  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.39% Indels: 0  
 DB: 8 Gaps: 0

US-10-047-825-4 (1-313) x ARCPNC2 (1-1330)

OY 106 GUUAlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgasp 125  
 DB 539 GAGGCTGCTGCGCCGGTGTGTGTCGCGAGATATCTCGCCCTTGCTGCTGAC 598

RESULT 14  
 AF485268 1389 bp mRNA linear PIN 24-MAR-2002  
 AF485268 LOCUS  
 DEFINITION Gossypium hirsutum bacterial-induced galactol peroxidase (pod29)  
 mRNA, complete cds.  
 ACCESSION AF485268  
 VERSION AF485268.1 GI:19698451  
 KEYWORDS  
 SOURCE Gossypium hirsutum.  
 ORGANISM Gossypium hirsutum.

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.  
 1 (bases 1 to 1389)  
 Delannoy, E., Assigbetse, K., Jalloul, A., Marmey, P., Daniel, J.-F., Geiger, J.-P. and Nicole, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-FEB-2002) UR Resistance, IRD, 911, Avenue Agropolis, Montpellier 34032, France  
 FEATURES  
 SOURCE location/Qualifiers  
 1..1389 /organism="Gossypium hirsutum"  
 /culturvar="Reda B50"

gene  
 CDS  
 1..1389 /db\_xref="taxon:3635"  
 /tissue\_type="2 week old cotyledons"  
 /gene="pod29"  
 87..1073 /gene="pod29"  
 /EC\_number="1.11.1.7"  
 /product="bacterial-induced galactol peroxidase"  
 /codon\_start=1  
 /protein\_id="AA193154.1"  
 /db\_xref="GI:19698452"  
 /translation="MDVSCFSONYLVLTLLIANVLSVSGCTRGVGYRSCPRVSI  
 RVSTVSHFSDPTLAPLRLMHPDGFHGCDSITLIDGTEKTAEPNLRGLREY  
 IDDAKTOLEACPGVASCADILALAAKDSVLSGSLSMQVPRGRDVSQASDANL  
 EGRPSVDVQOKFAKGLNTQDITVLTGHTITGTAQCFRRLYFMTTGAGAPS  
 ITFAFVSQALACPQNGKANGANRVALDGTGSRPFSVSNLRNRGVLQDQALMN  
 STTFVQVRLGIRGLGTFNIEFGKSNMKNISGVTGIDGIRIKVSKVN"

BASE COUNT 388 a 282 c 319 g 400 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.8e-11 Length: 1389  
 Score: 20.00 Matches: 20  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.39% Indels: 0  
 DB: 8 Gaps: 0

US-10-047-825-4 (1-313) x AF485268 (1-1389)

OY 106 GUUAlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgasp 125  
 DB 417 GAGGCTGCTGCGCCGGTGTGTGTCGCGAGATATCTCGCCCTTGCTGCTGAT 476

RESULT 15

ATP02 1400 bp mRNA linear PIN 08-FEB-1999  
 LOCUS A.thaliana mRNA for peroxidase AtPp1a, clone EST 151A1577.  
 DEFINITION X99097.1 GI:1429224  
 ACCESSION X99097.1  
 VERSION X99097.1  
 KEYWORDS  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1400)  
 Wellinger, K.G., Jespersen, H.M., Kjaersgaard, J.V.H., Justesen, A.F., Oestergaard, L., Abelskov, A.K., Jensen, R.B., Hansen, L.N. and Rasmussen, S.K.

TITLE From expressed sequence tags to structure function evolution and expression of 28 ER-targeted Arabidopsis peroxidases  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1400)  
 Justesen, A.F.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-JUL-1996) A.F. Justesen, Institute of Molecular Biology, Department of Protein Chemistry, Oester Farimagsgade 2A, DK-1353 Copenhagen K, Denmark  
 3 (bases 1 to 1400)

REFERENCE  
 AUTHORS Justesen, A.F., Jespersen, H.M. and Wellinger, K.G.  
 TITLE Analysis of two incompletely spliced Arabidopsis cDNAs encoding novel types of peroxidase  
 JOURNAL Biochim. Biophys. Acta 1443 (1-2), 149-154 (1998)  
 MEDLINE 99057591  
 PUBMED 9838086

FEATURES  
 SOURCE location/Qualifiers  
 1..1400 /organism="Arabidopsis thaliana"  
 /culturvar="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="EST 151A1577"  
 /tissue\_type="stems"

/tissue\_type="leaf"  
 /tissue\_type="flowers"  
 /tissue\_type="root"  
 /tissue\_type="siliques"  
 /clone\_lib="lambda PR12"  
 join(44, 463, 573, .1139)  
 /note="ARP15a"  
 /codon\_start=1  
 /product="peroxisase"  
 /protein\_id="CA67551.1"  
 /db\_xref="GI:1483222"  
 /db\_xref="SPTREMBL:O43387"  
 /translation="MGLVRSLLCLITFLNCLIIISVHGQATRRPQVSGTRIGFYITTC  
 PRAETIVRNAVAGSSDPRIAPGILRMHFHDCFYQGGDSILISGANTERTAGPNLN  
 LOGFEVIDNAKTGLEACPEVVSADILALARDVILTQGTGMQVPTGRDGRVSLA  
 SNANNLPGFRDSVAVOQKFSALGNTFRLDVLVVGHTIGTAGGVFRRRLFNITGOT  
 ADPTIDPFLAQLQTCPPONGDSVRVDLDTGSGSTMDTSTYNNLSRGGLQSDQVL  
 WIDPATRPVVOQLMAPRSTENVEFARSWRMSNIGVYTGANGELIRVCSAVN"  
 .44, .463  
 /number=1  
 464, .572  
 /number=1  
 573, .1139  
 /number=2

exon  
 Intron  
 exon  
 BASE COUNT 404 a 297 g 386 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2,826-11 Length: 1400  
 Score: 20.00 Matches: 20  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6,398 Indels: 0  
 DB: 8 Gaps: 0

US-10-047-825-4 (1-313) x ATP02 (1-1400)

QY 106 GIUAlaAlAcYsProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAsp 125  
 Db 392 GAAGCCGCGTCTCCTGGAGTTGCTCTGTGCTGATATTATTAGCATTAGCCCGCTGCGAC 451

Search completed: May 3, 2003, 14:34:30  
 Job time : 2215 secs



GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 3, 2003, 13:15:33 ; Search time 190 seconds

(without alignments)  
3709,871 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 313

Sequence: 1 MASPTMOCIVANVSLSCVA.....GVLTGTGACQIRRCRVVNS 313

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Word size: 16

Total number of hits satisfying chosen parameters: 21

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODER=frame+ p2n model -DEV=yjh  
-Q/cgn2.1/USPTO.spool/US10047825/runat\_26042003\_112041\_7634/app\_query.fasta\_1.455  
-DB=N.Geneseq\_101002 -OPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-DOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdt  
-LIST=45 -DOCCALIGN=200 -THR\_SCORE=quality -THR\_MIN=16 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=plc -HEAPS-ext -HEAPS-ext=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10047825.ecgn.1.1.200.0runat\_26042003\_112041\_7634 -NCPU=6 -ICPU=3  
-NO\_XIPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMBOUR=120  
-MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N.Geneseq\_101002: \*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: \*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: \*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: \*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: \*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: \*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: \*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: \*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: \*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: \*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: \*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: \*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: \*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: \*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: \*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: \*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: \*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: \*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: \*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: \*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: \*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: \*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: \*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: \*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: \*

Printed: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	6.4	1041	21	AAC46841
2	20	6.4	1269	21	AAC48650
3	20	6.4	1271	21	AAC34570
4	17	5.4	364	21	AA68190
5	17	5.4	512	21	AA68156
6	17	5.4	586	21	AA68186
7	17	5.4	717	21	AA68114
8	17	5.4	1090	22	AAH45565
9	17	5.4	1106	21	AAC45055
10	17	5.4	1167	18	AAT66687
11	17	5.4	1167	20	AAV81423
12	17	5.4	1167	20	AAV69276
13	17	5.4	1191	22	AAH45561
14	17	5.4	1229	22	AAH45568
15	16	5.1	337	21	AA68177
16	16	5.1	1191	18	AAT66686
17	16	5.1	1191	20	AAV81422
18	16	5.1	1191	20	AAV69275
19	16	5.1	1218	21	AAC49694
20	16	5.1	1220	21	AAC36625
21	16	5.1	1433	22	AAH4090

## ALIGNMENTS

RESULT 1  
AAC46841  
ID AAC46841 standard; DNA; 1041 BP.  
AC AAC46841;  
AC 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 51602.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129645.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135253.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0137682.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 26-JUN-1999; 99US-0140635.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142900.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143562.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145921.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151138.  
PR 01-SEP-1999; 99US-0151130.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 08-OCT-1999; 99US-0158369.  
PR 12-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.



PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145226.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147312.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151060.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154038.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159325.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160761.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Alignment Scores:  
Pred. No.: 9.13e-10  
Score: 20.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 6.39%  
DB: 21  
Gaps: 0  
  
US-10-047-825-4 (1-313) x AAC48650 (1-1269)  
OY 106 GLUAlaAlaCysProGlyValIValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAsp 125  
DB 403 GAAGCCGCGTGTCTGAGAGTGTCTGTCTGTATATTTACATATAGCCGCTGTAC 462  
  
RESULT 3  
AAC34570  
ID AAC34570 standard; DNA; 1271 BP.  
XX  
AC AAC34570;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7114.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134270.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142134.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158222.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160757.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 22-OCT-1999; 99US-0160999.

PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161923.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.:	9,14e-10	Length:	1271
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6,398	Indels:	0
DB:	21	Gaps:	0

US-10-047-825-4 (1-313) x AAC34570 (1-1271)

OY 106 GIuAlAlaCysPRGlyTAlValSerCysAlaAspIleuAlaLeuAlaAlaArgasp 125  
DB 405 GAAGCGCGTGTCTGTGAGTGTCTGTCTGTGATATTGAATTAAGCAATTAAGCGCTGTGAC 464

## RESULT 4

AAA68190 standard; DNA; 364 BP.

AAA68190;

24-OCT-2000 (first entry)

Pinus radiata peroxidase nucleotide sequence SEQ ID NO:366.

Pinus radiata; lignin biosynthetic pathway; Eucalyptus grandis;

Pinus radiata; Monterey pine; ds.

Pinus radiata.

WO200022099-A1.

20-APR-2000.

06-OCT-1999; 99WO-NZ00168.

09-OCT-1998; 98US-0169789.

14-JUL-1999; 99US-0143811.

(GENE-) GENESIS RES &amp; DEV CORP LTD.

(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Bjorsberg LN, Havukkala IJ;

WPI; 2000-317962/27.

Novel polynucleotide encoding enzymes involved in lignin-biosynthetic pathway useful for producing transgenic plants especially eucalyptus and pine species having altered lignin content, composition and structure -

Claim 1; Page 186; 213pp; English.

The present invention describes isolated polynucleotides and proteins encoding and representing the enzymes cinnamate 4-hydroxylase (C4H), coumarate 3-hydroxylase (C3H), phenolase (PVL), O-methyl transferase (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL), coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG), laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase, caffeic acid methyl transferase, caffeoyl CoA methyl transferase, coumarate CoA ligase, cytochrome P450 LXXII, diphenol oxidase, flavanol glucosyl transferase, flavanoid hydroxylase, and isoflavone reductase,

CC which are involved in the lignin biosynthetic pathway. The  
CC polynucleotides can be used for modulating lignin content, lignin  
CC composition and the structure of a plant, especially eucalyptus and pine  
CC species, and for modifying the activity of an enzyme involved in lignin  
CC biosynthetic pathway, and for producing a plant having altered lignin  
CC content, composition and structure. They can be used for designing probes  
CC and primers useful for detecting similar DNA and RNA sequences in any  
CC organism and for PCR amplification. The lignin content can be efficiently  
CC modified using the polynucleotides. AAA67908 to AAA68201 and AAB16341 to  
CC AAB16449 represent polynucleotide and protein sequences used in the  
CC exemplification of the present invention.

Sequence 364 BP; 98 A; 111 C; 69 G; 86 T; 0 other;

Alignment Scores:			
Pred. No.:	2,45e-07	Length:	364
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5,43%	Indels:	0
DB:	21	Gaps:	0

US-10-047-825-4 (1-313) x AAA68190 (1-364)

OY 182 AlAluSerGlyAlaHisThrIleGlyGlnAlaArgCysThrPhearg 198  
DB 51 GCACTCTCGGAGCCTACATGAGCCAGCGGTGACACCAATTCAGA 101

## RESULT 5

AAA68156 standard; DNA; 512 BP.

AAA68156;

24-OCT-2000 (first entry)

Eucalyptus grandis peroxidase nucleotide sequence SEQ ID NO:249.

Eucalyptus grandis; lignin biosynthetic pathway; Eucalyptus grandis;

Pinus radiata; Monterey pine; ds.

Eucalyptus grandis.

WO200022099-A1.

20-APR-2000.

06-OCT-1999; 99WO-NZ00168.

09-OCT-1998; 98US-0169789.

14-JUL-1999; 99US-0143811.

(GENE-) GENESIS RES &amp; DEV CORP LTD.

(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Bjorsberg LN, Havukkala IJ;

WPI; 2000-317962/27.

Novel polynucleotide encoding enzymes involved in lignin-biosynthetic pathway useful for producing transgenic plants especially eucalyptus and pine species having altered lignin content, composition and structure -

Claim 1; Page 133; 213pp; English.

The present invention describes isolated polynucleotides and proteins encoding and representing the enzymes cinnamate 4-hydroxylase (C4H), coumarate 3-hydroxylase (C3H), phenolase (PVL), O-methyl transferase (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL), coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG), laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,

CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,  
CC coumarate CoA ligase, cytochrome P450 LXA1A, diphenol oxidase, flavanol  
CC glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,  
CC which are involved in the lignin biosynthetic pathway. The  
CC polynucleotides can be used for modulating lignin content, lignin  
CC composition and the structure of a plant, especially eucalyptus and pine  
CC species, and for modifying the activity of an enzyme involved in lignin  
CC biosynthetic pathway, and for producing a plant having altered lignin  
CC content, composition and structure. They can be used for designing probes  
CC and primers useful for detecting similar DNA and RNA sequences in any  
CC organism and for PCR amplification. The lignin content can be efficiently  
CC modified using the polynucleotides. AAA67908 to AAA68201 and AAB16341 to  
CC AAB16449 represent polynucleotide and protein sequences used in the  
CC exemplification of the present invention.

XX SQ Sequence 512 BP; 128 A; 124 C; 127 G; 133 T; 0 other;

#### Alignment Scores:

Pred. No.:	3,39e-07	Length:	512
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	21	Gaps:	0

US-10-047-825-4 (1-313) x AAA68156 (1-512)

OY 109 CysPrGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgasp 125

DB 362 TGCCCGCGGTTCATGTGGACATCTTGACACTACTCTAGAGAT 412

#### RESULT 6

AAA68186

ID AAA68186 standard; DNA; 586 BP.

XX, AAA68186;

AC 24-OCT-2000 (first entry)

DT Pinus radiata peroxidase nucleotide sequence SEQ ID NO:362.

XX DE Pinus radiata; lignin biosynthetic pathway; Eucalyptus grandis;

XX KM Pinus radiata; Monterey pine; ds.

XX OS Pinus radiata.

XX PN WO200022099-A1.

XX PD 20-APR-2000.

XX PF 06-OCT-1999; 99WO-NZ00168.

XX PR 09-OCT-1998; 98US-0169789.

XX PR 14-JUL-1999; 99US-0143811.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX P1 Bloksberg LN, Havukkala IU;

XX DR WPI; 2000-317962/27.

XX PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic

XX PT pathway useful for producing transgenic plants especially eucalyptus

XX PT and pine species having altered lignin content, composition and

XX PT structure

XX PS Claim 1; Page 187; 213pp; English.

XX CC The present invention describes isolated polynucleotides and proteins

XX CC encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),

XX CC coumarate 3-hydroxylase (C3H), phenolase (PNU), O-methyl transferase

XX CC (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase

CC (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (C4L),  
CC confierol glucosyl transferase (CCR), confierin beta-glucosidase (CBG),  
CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,  
CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,  
CC glucosyl transferase, cytochrome P450 LXA1A, diphenol oxidase, flavanol  
CC coumarate CoA ligase, cytochrome P450 LXA1A, diphenol oxidase, flavanol  
CC glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,  
CC which are involved in the lignin biosynthetic pathway. The  
CC polynucleotides can be used for modulating lignin content, lignin  
CC composition and the structure of a plant, especially eucalyptus and pine  
CC species, and for modifying the activity of an enzyme involved in lignin  
CC biosynthetic pathway, and for producing a plant having altered lignin  
CC content, composition and structure. They can be used for designing probes  
CC and primers useful for detecting similar DNA and RNA sequences in any  
CC organism and for PCR amplification. The lignin content can be efficiently  
CC modified using the polynucleotides. AAA67908 to AAA68201 and AAB16341 to  
CC AAB16449 represent polynucleotide and protein sequences used in the  
CC exemplification of the present invention.

XX SQ Sequence 586 BP; 184 A; 157 C; 101 G; 144 T; 0 other;

#### Alignment Scores:

Pred. No.:	3,86e-07	Length:	586
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	21	Gaps:	0

US-10-047-825-4 (1-313) x AAA68186 (1-586)

OY 182 AlaleuSerGlyAlaHisThrIleGlyAlaAlaArgCysThrThrPhearg 198

DB 91 GCACCTCAGTGTCTCATATGTCACAGCAGCGTCACACATTCAGA 141

#### RESULT 7

AAA68174

ID AAA68174 standard; DNA; 717 BP.

XX, AAA68174;

AC 24-OCT-2000 (first entry)

DT Eucalyptus grandis peroxidase nucleotide sequence SEQ ID NO:350.

XX DE Eucalyptus grandis; lignin biosynthetic pathway; Eucalyptus grandis;

XX KM Pinus radiata; Monterey pine; ds.

XX OS Eucalyptus grandis.

XX PN WO200022099-A1.

XX PD 20-APR-2000.

XX PF 06-OCT-1999; 99WO-NZ00168.

XX PR 09-OCT-1998; 98US-0169789.

XX PR 14-JUL-1999; 99US-0143811.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX P1 Bloksberg LN, Havukkala IU;

XX DR WPI; 2000-317962/27.

XX PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic

XX PT pathway useful for producing transgenic plants especially eucalyptus

XX PT and pine species having altered lignin content, composition and

XX PT structure

XX PS Claim 1; Page 183; 213pp; English.

XX CC The present invention describes isolated polynucleotides and proteins

CC encoding and representing the enzymes cinaminate 4-hydroxylase (C4H),  
CC coumarate 3-hydroxylase (C3H), phenolase (PUL), O-methyl transferase  
CC (OMT), cinamyl alcohol dehydrogenase (CAD), cinamoyl-CoA reductase  
CC (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),  
CC confertifol glucosyl transferase (CGR), confertifol beta-glucosidase (CBG),  
CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,  
CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,  
CC coumarate CoA ligase, cytochrome P450 1XX1A, diphenol oxidase, flavanol  
CC glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase,  
CC which are involved in the lignin biosynthetic pathway. The  
CC polynucleotides can be used for modulating lignin content, lignin  
CC composition and the structure of a plant, especially eucalyptus and pine  
CC species, and for modifying the activity of an enzyme involved in lignin  
CC biosynthetic pathway, and for producing a plant having altered lignin  
CC content, composition and structure. They can be used for designing probes  
CC and primers useful for detecting similar DNA and RNA sequences in any  
CC organism and for PCR amplification. The lignin content can be efficiently  
CC modified using the polynucleotides. AAC67908 to AAC68201 and AAC6341 to  
CC AAC6449 represent polynucleotide and protein sequences used in the  
CC exemplification of the present invention.

XX Sequence 717 BP; 182 A; 191 C; 181 G; 163 T; 0 other;

#### Alignment Scores:

Pred. No.:	4.67e-07	Length:	717
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	21	Gaps:	0

US-10-047-825-4 (1-313) x AAC68174 (1-717)

OY 109 CysProGlyValValSerCysAlaAspIleuAlaLeuAlaAlaArgAsp 125  
|||||  
DB 513 TGCCTGGAGTGTCTCTTGGCCGACATCTTGCCCTTGCCGCGCGGAT 563

#### RESULT 8

AAC45565  
ID AAC45565 standard; DNA; 1090 BP.

AC AAC45565;

DT 13-SEP-2001 (first entry)

DE Peroxidase related DNA fragment SEQ ID 14.

XX Peroxidase; active oxygen resistance; transgenic plant; ds;  
KW environmental stress.

OS Nicotiana tabacum.

PN JP2001095585-A.

PD 10-APR-2001.

PR 30-SEP-1999; 99JP-0279690.

PR 30-SEP-1999; 99JP-0279690.

PA (TOYOTA JIDOSHA KK.

DR WPI: 2001-360494/38.

XX Peroxidase derived from Parquat-resistant callus, and gene encoding  
PT it, used for the development of plants resistant to active oxygen  
PT formed under environmental stress -

XX Example 4; Page 18; 23pp; Japanese.

XX This invention relates to a peroxidase derived from a Parquat resistant  
CC callus. Included in the invention are the gene encoding the peroxidase,  
CC a vector containing the gene, and a method for the preparation of the

CC peroxidase. The gene is useful for the development of a plant highly  
CC resistant against active oxygen which is formed under various  
CC environmental stress conditions. This sequence represents a peroxidase  
CC related DNA fragment.

XX Sequence 1090 BP; 280 A; 224 C; 259 G; 327 T; 0 other;

#### Alignment Scores:

Pred. No.:	6.95e-07	Length:	1090
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	22	Gaps:	0

US-10-047-825-4 (1-313) x AAC45565 (1-1090)

OY 109 CysProGlyValValSerCysAlaAspIleuAlaLeuAlaAlaArgAsp 125  
|||||  
DB 408 TGTCTGGAGTGTCTCTTGGCGATATCTTGTCTACTCTCTGTAT 458

#### RESULT 9

AAC45025  
ID AAC45025 standard; DNA; 1106 BP.

AC AAC45025;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 45019.

XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;

XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PR 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132486.

PR 08-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132487.

PR 14-MAY-1999; 99US-0132487.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.



PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140993.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141847.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144637.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159639.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.



CC amino acid identity, the coding regions of SEPB1 and SEPB2 share  
CC 98% amino acid identity, and the coding regions of SEPA1 and SEPB1  
CC share 47% amino acid identity. The peroxidases are useful in pulp  
CC and paper bleaching, on-site waste destruction, soil remediation,  
CC organic syntheses and diagnostic chemistries. Soybean peroxidase  
CC has advantages over chlorine bleach, being cheaper, more  
CC environmentally friendly, and producing hydroxyl ions with twice  
CC the oxidizing power of chlorine ions. The plant enzyme is cheap  
CC and easy to produce. The invention also relates to immunoassays or  
CC oligonucleotide assays which utilize soybean peroxidase as marker.

XX Sequence 1167 BP; 385 A; 232 C; 231 G; 319 T; 0 other;  
SQ

## Alignment Scores:

Pred. No.:	7,41e-07	Length:	1167
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	20	Gaps:	0

US-10-047-825-4 (1-313) x AAV69276 (1-1167)

OY 109 CysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAArgasp 125  
DB 369 TGCCAGAGTGTGCTCTTGTCTGACATCCTTGTCTAGCAGCAAGGAT 419

## RESULT 12

AAV69276  
ID AAV69276 standard; cDNA; 1167 BP.

AC AAV69276;

DT 29-JAN-1999 (first entry)

XX Soybean peroxidase SEPB2 polypeptide encoding cDNA.

XX Soybean peroxidase; SEPA1; SEPA2; SEPB1; SEPB2; recombinant;

KW Industry; diagnostic chemistry; ss.

OS Glycine max.

XX Key Location/Qualifiers

FT 5'UTR 1..38

FT CDS /tag= a

FT /tag= b

FT sig-peptide 39..101

FT mat-peptide 102..977

FT 3'UTR 978..1167

XX US5840558-A.

XX 24-NOV-1998.

XX 27-OCT-1995; 95US-0671320.

XX 27-OCT-1995; 95US-0671320.

XX (INDI-) INDIANA CROP IMPROVEMENT ASSOC.

XX Vierling RA;

XX WPI, 1999-034041/03.

XX P-PSDB; AAW81496.

XX cDNA encoding soya peroxidase SEPA1 - useful for producing

XX recombinant peroxidase

PS Example 7; Columns 35-38; 31pp; English.

XX This cDNA encodes a soybean peroxidase SEPB2 polypeptide. The invention  
CC provides four cDNA sequences (AAV69273 to AAV69276) encoding soybean  
CC peroxidases SEPA1, SEPA2, SEPB1 and SEPB2 (AAW81493 to AAW81496)  
CC respectively. An expression vector containing the SEPA1 cDNA sequence  
CC can be used to transform host cells for the recombinant production of  
CC the peroxidase. The peroxidase is useful in industries and diagnostic  
CC chemistries.

XX Sequence 1167 BP; 385 A; 232 C; 231 G; 319 T; 0 other;  
SQ

## Alignment Scores:

Pred. No.:	7,41e-07	Length:	1167
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	20	Gaps:	0

US-10-047-825-4 (1-313) x AAV69276 (1-1167)

OY 109 CysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAArgasp 125  
DB 369 TGCCAGAGTGTGCTCTTGTCTGACATCCTTGTCTAGCAGCAAGGAT 419

## RESULT 13

AAH45561  
ID AAH45561 standard; DNA; 1191 BP.

AC AAH45561;

DT 13-SEP-2001 (first entry)

XX Peroxidase gene.

XX Peroxidase gene.

KW Peroxidase; active oxygen resistance; transgenic plant;

XX environmental stress; ds.

OS Nicotiana tabacum.

XX Key Location/Qualifiers

FT CDS 49..1041

FT /tag= a

FT /product= "Peroxidase"

PN JP2001095585-A.

XX 10-APR-2001.

XX 30-SEP-1999; 99JP-0279690.

XX 30-SEP-1999; 99JP-0279690.

XX (TOYT ) TOYOTA JIDOSHA KK.

XX WPI; 2001-360494/38.

XX P-PSDB; AAG64621.

XX Peroxidase derived from Paraquat-resistant callus, and gene encoding

XX it, used for the development of plants resistant to active oxygen

XX formed under environmental stress -

XX Claim 3; Page 13-15; 23pp; Japanese.

XX This invention relates to a peroxidase derived from a Paraquat resistant

XX callus. Included in the invention are the gene encoding the peroxidase,

XX a vector containing the gene, and a method for the preparation of the

XX peroxidase. The gene is useful for the development of a plant highly

XX resistant against active oxygen which is formed under various

XX environmental stress conditions. The present sequence represents the

XX peroxidase gene of the invention.

SQ Sequence 1191 BP; 326 A; 245 C; 277 G; 343 T; 0 other;

# Alignment Scores:

Pred. No.:	7.56e-07	Length:	1191
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	22	Gaps:	0

US-10-047-825-4 (1-313) x AAH45561 (1-1191)

OY 109 CysPProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgasp 125

DB 397 TGCCGTGAGATTGTTCTGCTGACATCTTGCTGCTGCTGCTGAT 447

# RESULT 14

AAH45568 standard; cDNA; 1229 BP.

AAH45568;

13-SEP-2001 (first entry)

PER4 peroxidase related cDNA SEQ ID 17.

Peroxidase; active oxygen resistance; transgenic plant;

environmental stress; ss.

Nicotiana tabacum.

JP2001095585-A.

10-APR-2001.

30-SEP-1999; 99JP-0279690.

30-SEP-1999; 99JP-0279690.

(TOYOTA) TOYOTA JIDOSHA KK.

WPI; 2001-360494/38.

Peroxidase derived from Paraquat-resistant callus, and gene encoding

it, used for the development of plants resistant to active oxygen

formed under environmental stress -

Example 5; Page 20-21; 23pp; Japanese.

This invention relates to a peroxidase derived from a Paraquat resistant

callus. Included in the invention are the gene encoding the peroxidase,

a vector containing the gene, and a method for the preparation of the

peroxidase. The gene is useful for the development of a plant highly

resistant against active oxygen which is formed under various

environmental stress conditions. This sequence represents a PER4

peroxidase related cDNA.

Sequence 1229 BP; 364 A; 238 C; 273 G; 354 T; 0 other;

# Alignment Scores:

Pred. No.:	7.79e-07	Length:	1229
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	22	Gaps:	0

US-10-047-825-4 (1-313) x AAH45568 (1-1229)

OY 109 CysPProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgasp 125

DB 408 TGCCGTGAGATTGTTCTGCTGACATCTTGCTGCTGCTGAT 458

# RESULT 15

AAA68177 standard; DNA; 337 BP.

AAA68177;

24-OCT-2000 (first entry)

Eucalyptus grandis peroxidase nucleotide sequence SEQ ID NO:353.

Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;

Pinus radiata; Monterey pine; ds.

Eucalyptus grandis.

WO20002209-A1.

20-APR-2000.

06-OCT-1999; 99WO-N200168.

09-OCT-1998; 98US-0169789.

14-JUL-1999; 99US-0143811.

(GENE-) GENESIS RES & DEV CORP LTD.

(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Bloksberg LN, Havukkala IJ;

WPI; 2000-317962/27.

Novel polynucleotide encoding enzymes involved in lignin-biosynthetic

pathway useful for producing transgenic plants especially eucalyptus

and pine species having altered lignin content, composition and

structure -

Claim 1; Page 184; 213pp; English.

The present invention describes isolated polynucleotides and proteins

encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),

coumarate 3-hydroxylase (C3H), phenolase (PNU), 0-methyl transferase

(OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase

(CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),

coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),

laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,

caffeic acid methyl transferase, caffeoyl CoA methyl transferase,

coumarate CoA ligase, cytochrome P450 1X1A, diphenol oxidase, flavanol

glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,

which are involved in the lignin biosynthetic pathway. The

polynucleotides can be used for modulating lignin content, lignin

composition and the structure of a plant, especially eucalyptus and pine

species, and for modifying the activity of an enzyme involved in lignin

biosynthetic pathway, and for producing a plant having altered lignin

content, composition and structure. They can be used for designing probes

and primers useful for detecting similar DNA and RNA sequences in any

organism and for PCR amplification. The lignin content can be efficiently

modified using the polynucleotides. AAA67908 to AAA68201 and AAA6341 to

AAA6449 represent polynucleotide and protein sequences used in the

exemplification of the present invention.

Sequence 337 BP; 74 A; 90 C; 85 G; 88 T; 0 other;

# Alignment Scores:

Pred. No.:	2.19e-06	Length:	337
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.11%	Indels:	0
DB:	21	Gaps:	0

US-10-047-825-4 (1-313) x AAA68177 (1-337)

OY 53 ArgMetGlyAlaSerLeuLeuArgLeuPheHisaspCysPheVal 68

Wed May 7 14:10:02 2003

us-10-047-825-4.rng

Page 13

Db 173 CGCATGGGTCTTCTCTCTGGCCCTCTTCTTCATGATGCTTCGTC 220

```
Search completed: May 3, 2003, 13:57:21
Job time : 194 secs
```



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 3, 2003, 13:51:59 ; Search time 46 Seconds  
(without alignments)  
2086.735 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 313  
Sequence: 1 MASPTLMQCLVAVSLSCVA.....VGVLTRAGQIRRCRVNS 313

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 15338381 residues

Word size: 16

Total number of hits satisfying chosen parameters: 9

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10047825/runal.26042003\_112043\_7672/app\_query.fasta\_1.455  
-DB=Issued\_Patents.NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=16 -ALIGN=15 -MODS=LOCAL  
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10047825 -GCN1\_1\_32\_0nauc.26042003\_112043\_7672 -NCPUR=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAB -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV\_TIMEROUT=120  
-WARN\_TIMEROUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents.NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	17	5.4	364	4	US-09-615-192A-366
2	17	5.4	512	4	US-09-615-192A-249
3	17	5.4	586	4	US-09-615-192A-362
4	17	5.4	717	4	US-09-615-192A-350
5	17	5.4	1167	2	US-08-671-320-16
6	17	5.4	1167	2	US-08-671-320-16
7	16	5.1	337	4	US-09-615-192A-353
8	16	5.1	1191	2	US-08-671-320-14
9	16	5.1	1191	2	US-08-668-577-14

#### ALIGNMENTS

RESULT 1  
US-09-615-192A-366  
; Sequence 366, Application US/09615192A  
; Patent No. 6410718  
; GENERAL INFORMATION:  
; APPLICANT: Blosberg, Leonard N.  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1003c40  
; CURRENT APPLICATION NUMBER: US/09/615,192A  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 08/975,316  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: US 08/713,000  
; PRIOR FILING DATE: 1996-09-11  
; PRIOR APPLICATION NUMBER: US 09/169,789  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 405  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 366  
; LENGTH: 364  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-615-192A-366

Alignment Scores:  
Pred. No.: 1.92e-08 Length: 364  
Score: 17.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.43% Indels: 0  
Gaps: 0  
DB: 4

US-10-047-825-4 (1-313) x US-09-615-192A-366 (1-364)

QY 182 AlaleuserglyalshlrhrleglyglnAlaIrgCysThrThphea9 198  
|||||  
Db 51 GCACCTCGGAGCTCATGACATTGGCCAGCGCGTGACACCATTCAGA 101

RESULT 2  
US-09-615-192A-249  
; Sequence 249, Application US/09615192A  
; Patent No. 6410718  
; GENERAL INFORMATION:  
; APPLICANT: Blosberg, Leonard N.  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1003c40  
; CURRENT APPLICATION NUMBER: US/09/615,192A  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 08/975,316  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: US 08/713,000  
; PRIOR FILING DATE: 1996-09-11  
; PRIOR APPLICATION NUMBER: US 09/169,789  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 405  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 249  
; LENGTH: 512  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-09-615-192A-249

Alignment Scores:  
Pred. No.: 2.67e-08 Length: 512  
Score: 17.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.43% Indels: 0

DB: 4 Gaps: 0

US-10-047-825-4 (1-313) x US-09-615-192A-249 (1-512)

QY 109 CysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgASP 125  
|||||  
DB 362 TGCCCGGGGCTGCTTCATGTCGACACATCTTGCGCATGCTGCTAGAGAT 412

RESULT 3

US-09-615-192A-362

; Sequence 362, Application US/09615192A

; Patent No. 6410718

; GENERAL INFORMATION:

; APPLICANT: Bloksberg, Leonard N.

; APPLICANT: Havukkala, Ilkka

; TITLE OF INVENTION: Materials and Methods for the

; FILE REFERENCE: 11000.1003c4U

; CURRENT APPLICATION NUMBER: US/09/615,192A

; CURRENT FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 08/975,316

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: US 08/713,000

; PRIOR FILING DATE: 1996-09-11

; PRIOR APPLICATION NUMBER: US 09/169,789

; PRIOR FILING DATE: 1998-10-09

; NUMBER OF SEQ ID NOS: 405

; SOFTWARE: FastSeq for Windows, Version 3.0

; SEQ ID NO 362

; LENGTH: 586

; TYPE: DNA

; ORGANISM: Pinus radiata

US-09-615-192A-362

Alignment Scores:

Pred. No.: 3,05e-08 Length: 586

Score: 17.00 Matches: 17

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.43% Indels: 0

DB: 4 Gaps: 0

US-10-047-825-4 (1-313) x US-09-615-192A-362 (1-586)

QY 182 AldeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrThPhaarg 198  
|||||  
DB 91 GCACCTCTCAGGTGCTCATCAATTGCTCAAGCAGCGTGACCATTCAGA 141

RESULT 4

US-09-615-192A-350

; Sequence 350, Application US/09615192A

; Patent No. 6410718

; GENERAL INFORMATION:

; APPLICANT: Bloksberg, Leonard N.

; APPLICANT: Havukkala, Ilkka

; TITLE OF INVENTION: Materials and Methods for the

; FILE REFERENCE: 11000.1003c4U

; CURRENT APPLICATION NUMBER: US/09/615,192A

; CURRENT FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 08/975,316

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: US 08/713,000

; PRIOR FILING DATE: 1996-09-11

; PRIOR APPLICATION NUMBER: US 09/169,789

; PRIOR FILING DATE: 1998-10-09

; NUMBER OF SEQ ID NOS: 405

; SOFTWARE: FastSeq for Windows, Version 3.0

; SEQ ID NO 350

; LENGTH: 717

; TYPE: DNA

; ORGANISM: Eucalyptus grandis

US-09-615-192A-350

Alignment Scores:

Pred. No.: 3.71e-08 Length: 717

Score: 17.00 Matches: 17

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.43% Indels: 0

DB: 4 Gaps: 0

US-10-047-825-4 (1-313) x US-09-615-192A-350 (1-717)

QY 109 CysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgASP 125  
|||||  
DB 513 TGCCCGGGGCTGCTTCATGTCGACACATCTTGCGCATGCTGCTAGAGAT 563

RESULT 5

US-08-671-320-16

; Sequence 16, Application US/08671320

; Patent No. 5840558

; GENERAL INFORMATION:

; APPLICANT: VIERLING JR, RICHARD A

; TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN

; FILE REFERENCE: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESS: 1125 SO. 103RD STREET

; STREET: SUITE 330

; CITY: OMAHA

; STATE: NE

; COUNTRY: US

; ZIP: 68124-1076

; COMPUTER READABLE FORM:

; MEDIUM TYPE: floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA: US/08/671,320

; APPLICATION NUMBER: US/08/671,320

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: JONDLER, ROBERT J.

; REGISTRATION NUMBER: 33,915

; REFERENCE/DOCKET NUMBER: 1227-001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 402-398-9005

; TELEFAX: 402-398-9005

; INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 1167 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: 5'UTR

LOCATION: 1..38

FEATURE:

NAME/KEY: CDS

LOCATION: 39..977

FEATURE:

NAME/KEY: 3'UTR

LOCATION: 978..1167

FEATURE:

NAME/KEY: sig\_peptide

LOCATION: 39..101

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 102..977

US-08-671-320-16

Alignment Scores:

Pred. No.: 5.96e-08 Length: 1167



Score: 17.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.43% Indels: 0  
DB: 2 Gaps: 0

US-10-047-825-4 (1-313) x US-08-671-320-16 (1-1167)

QY 109 CysProGlyValIserCysAlaAspIleLeuAlaLeuAlaAlaArgasp 125  
DB 369 TGCACGGTGCTGCTCTTGTGCTGACATCTTGTCTAGCAGCAGGAT 419

RESULT 6

US-08-868-577-16  
; Sequence 16, Application US/08868577  
; Patent No. 5866695

GENERAL INFORMATION:

APPLICANT: Vierling Jr., Richard A  
TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Flg9, Ernst & Kurz  
STREET: 555 13th Street NW, Suite 701 East  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/868,577  
FILING DATE: 04-JUN-1997

CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Jondle, Robert J.  
REGISTRATION NUMBER: 33,915  
REFERENCE/DOCKET NUMBER: N1227-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 402-333-1550  
TELEFAX: 402-333-1510

INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1167 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..38

FEATURE:  
NAME/KEY: CDS  
LOCATION: 39..977

FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 978..1167

FEATURE:  
NAME/KEY: s1g-peptide  
LOCATION: 39..101

FEATURE:  
NAME/KEY: mat-peptide  
LOCATION: 102..977

US-08-868-577-16

Alignment Scores:

Pred. No.: 5.96e-08  
Score: 17.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%

Length: 1167  
Matches: 17  
Conservative: 0  
Mismatches: 0

Query Match: 5.43% Indels: 0  
DB: 2 Gaps: 0

US-10-047-825-4 (1-313) x US-08-868-577-16 (1-1167)

QY 109 CysProGlyValIserCysAlaAspIleLeuAlaLeuAlaAlaArgasp 125  
DB 369 TGCACGGTGCTGCTCTTGTGCTGACATCTTGTCTAGCAGCAGGAT 419

RESULT 7

US-09-615-192A-353  
; Sequence 353, Application US/09615192A  
; Patent No. 6410718

GENERAL INFORMATION:

APPLICANT: Bloksberg, Leonard N.  
TITLE OF INVENTION: Materials and Methods for the  
Modification of Plant Lignin Content  
FILE REFERENCE: 11000.1003C4U  
CURRENT FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 08/975,316  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11  
PRIOR APPLICATION NUMBER: US 09/169,789  
NUMBER OF SEQ ID NOS: 405  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 353  
LENGTH: 337  
TYPE: DNA

ORGANISM: Eucalyptus grandis  
US-09-615-192A-353

Alignment Scores:

Pred. No.: 1.83e-07 Length: 337  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.11% Indels: 0  
DB: 4 Gaps: 0

US-10-047-825-4 (1-313) x US-09-615-192A-353 (1-337)

QY 53 ArgMetGlyAlaSerLeuLeuArgLeuPheHisAspCysPheVal 63  
DB 173 CGCATGGGTGCTCTTGTGCTGACATCTTGTCTAGCAGCAGGAT 220

RESULT 8

US-08-671-320-14  
; Sequence 14, Application US/08671320  
; Patent No. 5840558

GENERAL INFORMATION:

APPLICANT: Vierling Jr., Richard A  
TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: 1125 SO. 103RD STREET  
STREET: SUITE 330  
CITY: OMAHA  
STATE: NE  
COUNTRY: US  
ZIP: 68124-1076

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/671,320  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: JONDLE, ROBERT J.  
REGISTRATION NUMBER: 33,915  
REFERENCE/DOCKET NUMBER: 1227-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 402-398-9000  
TELEFAX: 402-398-9005  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1191 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..59  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 60..998  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 999..1191  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 60..122  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 123..998  
US-08-671-320-14  
Alignment Scores:  
Pred. No.: 6.25e-07 Length: 1191  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.11% Indels: 0  
DB: 2 Gaps: 0  
US-10-047-825-4 (1-313) x US-08-671-320-14 (1-1191)  
Qy 109 CysProGlyValValSerCysAlaSpIleLeuAlaLeuAlaAlaarg 124  
Db 390 TGCCCTGGTGTGTCTCTGTGCTGACATCCTCTGCGCAGCAAG 437  
RESULT 9  
US-08-868-577-14  
Sequence 14, Application US/08868577  
Patent No. 5866695  
GENERAL INFORMATION:  
APPLICANT: Vierling Jr., Richard A  
TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurtz  
STREET: 555 13th Street NW, Suite 701 East  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/868,577  
FILING DATE: 04-JUN-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Jondle, Robert J.

REGISTRATION NUMBER: 33,915  
REFERENCE/DOCKET NUMBER: N1227-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 402-333-1550  
TELEFAX: 402-333-1510  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1191 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 60..998  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..59  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 999..1191  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 60..122  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 123..998  
US-08-868-577-14  
Alignment Scores:  
Pred. No.: 6.25e-07 Length: 1191  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.11% Indels: 0  
DB: 2 Gaps: 0  
US-10-047-825-4 (1-313) x US-08-868-577-14 (1-1191)  
Qy 109 CysProGlyValValSerCysAlaSpIleLeuAlaLeuAlaAlaarg 124  
Db 390 TGCCCTGGTGTGTCTCTGTGCTGACATCCTCTGCGCAGCAAG 437  
Search completed: May 3, 2003, 14:53:14  
Job time: 48 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 3, 2003, 14:34:34 ; Search time 108 Seconds

(without alignments)  
3424.513 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 313

Sequence: 1 MASPTLMQCLVAVSLSCVA.....VGVLTGATGATRCRCRVVNS 313

## Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 746064 seqs, 590810554 residues

Word size: 16

Total number of hits satisfying chosen parameters: 7

Minimum DB seg length: 0  
Maximum DB seg length: 2000000000

Post-processing: Listing first 45 summaries

## Command line parameters:

-MODEL=frame+\_p2n.model -DEV-xlh  
-Q=/cgn2.1/USPTO/US10047825/rnatc\_26042003.112044\_7756/app.query.fasta\_1.455  
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rrpb -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR.SCORE=quality THR\_MIN=16  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10047825 -cgn\_1.1\_84\_etunal\_26042003.112044\_7756  
-NCP=6 -ICP=3 -NO\_XLPTX -NO\_MMAP -LAKEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEDOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

## Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PTC\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/PTC\_NEW\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	6.4	960	US-09-938-842A-1596	Sequence 1596, Ap
2	20	6.4	987	US-09-938-842A-1518	Sequence 1518, Ap
3	17	5.4	271	US-09-878-574-12464	Sequence 12464, A
4	17	5.4	334	US-09-878-574-744	Sequence 744, App

5	17	5.4	993	US-09-938-842A-2693	Sequence 2693, Ap
6	17	5.4	1074	US-09-938-842A-2656	Sequence 2656, Ap
7	16	5.1	285	US-09-878-574-8092	Sequence 8092, Ap

## ALIGNMENTS

```

RESULT 1
US-09-938-842A-1596
; Sequence 1596, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1518
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

Alignment Scores:
Pred. No.: 5.5e-11
Score: 20.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 6.39%
DB: 9
Gaps: 0

US-10-047-825-4 (1-313) x US-09-938-842A-1596 (1-960)
QY 106 GUUAIAAAGCSPGCIYVAIVAIserCysAlaAspI]eleuAlaLeuAlaIlaArgsp 125
DB 319 GAGCGACGATGCTCGGTGTGTCTCTTGTCTGATATCTTGCCCTTACGCGCTGTGTAT 378

RESULT 2
US-09-938-842A-1518
; Sequence 1518, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1518
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

```

US-09-938-842A-1518

## Alignment Scores:

Pred. No.:	5.69e-11	Length:	987
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.39%	Indels:	0
DB:	9	Gaps:	0

US-10-047-825-4 (1-313) x US-09-938-842A-1518 (1-987)

QY 106 GLUAlaAlaCysProGlyValIValSerCysAlaAspIleuAlaIleuAlaAlaArgasp 125

DB 349 GAAGCGCGGTCTCCGAGATGTCTCTTGTCTGATATTTACGATAGCGCGTGTGAC 408

## RESULT 3

US-09-878-574-12464  
Sequence 12464, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:

APPLICANT: Byrum, Joseph R.

APPLICANT: La Rosa, Thomas J.

APPLICANT: Thompson, Michael D.

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

FILE REFERENCE: 38-21(13401)B

CURRENT APPLICATION NUMBER: US/09/878,574

CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 09/333,535

PRIOR FILING DATE: 1999-06-14

NUMBER OF SEQ ID NOS: 15775

SEQ ID NO 12464

LENGTH: 271

TYPE: DNA

ORGANISM: Glycine max

OTHER INFORMATION: Clone ID: 701065809H1

US-09-878-574-12464

## Alignment Scores:

Pred. No.:	1.95e-08	Length:	271
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	10	Gaps:	0

US-10-047-825-4 (1-313) x US-09-878-574-12464 (1-271)

QY 60 ArgLeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleu 76

DB 220 CGTCTCTTCTTCATGACTGCTTGTCTTCACGCTGTGATGCGTCACTACTA 270

## RESULT 4

US-09-878-574-744  
Sequence 744, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:

APPLICANT: Byrum, Joseph R.

APPLICANT: La Rosa, Thomas J.

APPLICANT: Thompson, Michael D.

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

FILE REFERENCE: 38-21(13401)B

CURRENT APPLICATION NUMBER: US/09/878,574

CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 09/333,535

PRIOR FILING DATE: 1999-06-14

NUMBER OF SEQ ID NOS: 15775

SEQ ID NO 744

LENGTH: 334

TYPE: DNA

ORGANISM: Glycine max

OTHER INFORMATION: Clone ID: LIB3028-046-Q1-B1-G2

US-09-878-574-744

Alignment Scores:			
Pred. No.:	2.35e-08	Length:	334
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	10	Gaps:	0

US-10-047-825-4 (1-313) x US-09-878-574-744 (1-334)

QY 62 PhePheHisAspCysPheValGlnGlyCysAspGlySerIleuLeuasp 78

DB 249 TCTCTCCATGACTGCTTGTCTTCACGAGTTCGATGCGTGTGATGTGGAT 299

## RESULT 5

US-09-938-842A-2693  
Sequence 2693, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

FILE REFERENCE: SCRIPI300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 2693

LENGTH: 993

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-2693

## Alignment Scores:

Pred. No.:	6.24e-08	Length:	993
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.43%	Indels:	0
DB:	9	Gaps:	0

US-10-047-825-4 (1-313) x US-09-938-842A-2693 (1-993)

QY 109 CysProGlyValIValSerCysAlaAspIleuAlaIleuAlaAlaArgasp 125

DB 361 TGTCCGAGATGTGTGATGCGGATATACCTCTGCGCGCTAGAGAC 411

## RESULT 6

US-09-938-842A-2656  
Sequence 2656, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPI300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 26, 2003, 12:31:33 ; Search time 46 Seconds  
(without alignments)  
654.132 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 1604

Sequence: 1 MASPTMOCIVANVSLSCVA.....VGVLGTAGQIRRCRVNS 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*\*\*  
2: PIR1:\*\*\*  
3: PIR2:\*\*\*  
4: PIR3:\*\*\*  
5: PIR4:\*\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1004.5	62.6	309	2 T09166	probable peroxidase (EC 1.1
2	997.5	62.2	305	2 T08121	peroxidase (EC 1.1
3	972.5	60.6	323	2 T09164	probable peroxidase (EC 1.1
4	942	58.7	296	1 OPNB7	peroxidase (EC 1.1
5	935	58.3	315	2 T10252	peroxidase (EC 1.1
6	928	57.9	317	2 S22087	peroxidase (EC 1.1
7	926	57.7	314	2 T03929	peroxidase (EC 1.1
8	923	57.5	316	2 A38265	peroxidase (EC 1.1
9	915.5	57.1	316	2 S61405	peroxidase (EC 1.1
10	895.5	55.8	319	2 S61408	peroxidase (EC 1.1
11	894	55.7	322	2 B65555	peroxidase (EC 1.1
12	867	54.1	314	2 S61406	peroxidase (EC 1.1
13	849.5	53.0	312	2 S13325	peroxidase (EC 1.1
14	846.5	52.8	315	2 T09165	peroxidase (EC 1.1
15	838.5	52.3	312	2 S13375	probable peroxidase (EC 1.1
16	818	51.0	315	2 T06164	peroxidase (EC 1.1
17	812	50.6	315	2 T06172	peroxidase (EC 1.1
18	808	50.4	331	2 A85430	peroxidase (EC 1.1
19	801	49.9	338	2 H84560	probable peroxidase (EC 1.1
20	774	48.3	337	2 S60054	probable peroxidase (EC 1.1
21	770	48.0	347	2 S60054	probable peroxidase (EC 1.1
22	767.5	47.8	315	2 C86280	peroxidase (EC 1.1
23	762.5	47.5	288	2 S61407	peroxidase (EC 1.1
24	750	46.8	351	2 J04779	peroxidase (EC 1.1
25	747	46.6	343	2 T09566	peroxidase (EC 1.1
26	734	45.8	325	2 S32768	peroxidase (EC 1.1
27	732.5	45.7	343	2 T09565	peroxidase (EC 1.1
28	727.5	45.4	352	2 T05723	peroxidase (EC 1.1
29	723	45.1	324	2 A39889	peroxidase (EC 1.1

30	719	44.8	358	2 J04781	peroxidase (EC 1.1
31	718	44.8	327	2 S14268	peroxidase (EC 1.1
32	718	44.8	335	2 T03912	peroxidase (EC 1.1
33	715.5	44.6	336	2 T45730	peroxidase-like pr
34	715	44.6	322	2 T03686	peroxidase (EC 1.1
35	713.5	44.5	332	2 S00627	peroxidase (EC 1.1
36	708	44.1	355	2 J04780	peroxidase (EC 1.1
37	707	44.1	353	1 OPBHC	peroxidase (EC 1.1
38	706	44.0	350	2 C96532	peroxidase ATP5a l
39	702.5	43.8	351	2 S00626	peroxidase (EC 1.1
40	697.5	43.5	343	2 S60055	peroxidase (EC 1.1
41	695.5	43.4	336	2 H84767	probable peroxidase
42	692.5	43.2	353	2 T46118	peroxidase - Arabi
43	690	43.0	354	2 J00457	peroxidase (EC 1.1
44	687.5	42.9	318	2 J02217	peroxidase (EC 1.1
45	685.5	42.7	353	2 S37495	peroxidase (EC 1.1

## ALIGNMENTS

RESULT 1	
T09166	probable peroxidase (EC 1.11.1.7) (clone PC23) - spinach (fragment)
C:Species: Spinacia oleracea (spinach)	
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 20-Jun-2000	
C:Accession: T09166	
R:Simon, P.	
submitted to the EMBL Data Library, January 1997	
A:Reference number: Z16599	
A:Accession: T09166	
A:Status: preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: mRNA	
A:Residues: 1-309 <SIM>	
A:Cross-references: EMBL:Y10467	
A:Status: preliminary; translated from GB/EMBL/DBJ	
A:Experimental source: subspecies Nobel	
C:Genetics:	
A:Gene: prx16	
C:Superfamily: peroxidase	
C:Keywords: heme; iron; metalloprotein; oxidoreductase	
F:23-104/Disulfide bonds: #status predicted	
F:50/Active site: Arg #status predicted	
F:54.182/Binding site: heme iron (His) (axial ligands) #status predicted	
F:56-61/Disulfide bonds: #status predicted	
F:110-305/Disulfide bonds: #status predicted	
F:189-214/Disulfide bonds: #status predicted	
Query Match	62.6%; Score 1004.5; DB 2; Length 309;
Best Local Similarity	63.1%; Pred. No. 5.7e-74;
Matches 195; Conservative 46; Mismatches 61; Indels 7; Gaps 3;	
QY 11 VAVSLSCVAHQAOLSPFTFASCPNQLQIVRAAMQAVASERMAKASLLRFHHCFOQG 70	
DB 1 ILAVLACSNMQLSKHAKSPNLEKIVRTMVAQVOKERMAQSILRFFHHCFOQG 60	
QY 71 CDGSLD---AGEKTPAGPML-NSVRGFEVIDITIKKNVEACGVSACADIALAARD 125	
DB 61 CQASLLDPTSTFTGEKTAISNRNNSVGRFEVIDIKKNVEACATVSCADIALAARD 120	
QY 126 GNNLLGPTWSPVPLGRDSTASASLANSPPTPTASLGLTSLFGRGLSPRDMTALSG 185	
DB 121 GVFLLGPEWKPPLGRDARTASTLATNINLPPASSLSLNTLFFNNKGLSPKDMTALSG 180	
QY 186 ARTIGQARCTPFRGRTYGGTDINASFALRQOTGPRSG--GDGNTAPTDVQTPVRFDAY 243	
DB 181 ARTIGLARKVSRHRTIYNDTIDANFEATRKVNCPLSNNTGNTINLADLSPKRFDSY 240	
QY 244 FNNLSRRLGFHSQDELFPNGSDALVROYASASLFNADFEVAAMIRGNVGLTGAQ 303	
DB 241 YKNLAKRCLHSDQELVNGSGDALVTRYSKSNAPAKDFVAATIKKGNISPLTGSSE 300	
QY 304 IRRNCRVNA 312	

Db 301 IRKNCRFIN 309

## RESULT 2

T08121

peroxidase (EC 1.11.1.7) - flax (fragment)

C:Species: Linum usitatissimum (flax)

C:Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 04-Mar-2000

C:Accession: T08121

R:Omann, F.; Tyson, H.

submitted to the EMBL Data Library, February 1998

A:Description: cDNA sequence of a basic peroxidase (FLXPER4) in flax.

A:Reference number: 216366

A:Accession: T08121

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-305 <OMA>

A:Cross-references: EMBL:AF049881; NID:g2944416; PIDN:AAC05277.1; PID:g2944417

A:Experimental source: cv. Stormont Citrus

C:Genetics:

A:Gene: PER4

C:Superfamily: peroxidase

C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase

F:23-103/Disulfide bonds: #status predicted

F:50/Active site: Arg #status predicted

F:54,181/Binding site: heme iron (His) (axial ligands) #status predicted

F:56-61/Disulfide bonds: #status predicted

F:109-301/Disulfide bonds: #status predicted

F:188-213/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 62.2%; Score 997.5; DB 2; Length 305;

Matches 197; Conservative 37; Mismatches 66; Indels 5; Gaps 2;

12 ANSLSCVAHAQSLPFFYASCPNIOSTYRAMTQAVASEOMGMSLRPFHDFVGC 71

12 ANSLSCVAHAQSLPFFYASCPNIOSTYRAMTQAVASEOMGMSLRPFHDFVGC 71

2 ANSLSCVAHAQSLPFFYASCPNIOSTYRAMTQAVASEOMGMSLRPFHDFVGC 61

72 DGSIIID---AGGKTAGPNLNSVGFVIDITIKRNVPAACPGVSCADIALAARGT 127

62 DGSIIID---AGGKTAGPNLNSVGFVIDITIKRNVPAACPGVSCADIALAARGT 121

128 NLGGFTWVPLGRDSTTASASLANSPPTASLGLTISLFGROGLSPDMTALSGAH 187

122 VLVGGFTWVPLGRDSTTASASLANSPPTASLGLTISLFGROGLSPDMTALSGAH 181

188 TIGQACCTFRGRIYGDPTDINAFALRQTCPRSGGCGNLAPIDVOPVPRFDYFTNL 247

182 TIGQACCTFRGRIYGDPTDINAFALRQTCPRSGGCGNLAPIDVOPVPRFDYFTNL 240

248 LSRGLFHSODELFNGSODALVROYTSASLSFNADFYAAMIRMGVGLTGTACQIRRN 307

241 VARGRLHSDQELFNGTQDALVRYTSNNAATFATDFAAAMVAMGNIISLTGTNEIRN 300

QY 308 CRVYN 312

Db 301 CRRPN 305

## RESULT 3

T09164

probable peroxidase (EC 1.11.1.7) (clone PC44) - spinach

C:Species: Spinacia oleracea (spinach)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 20-Jun-2000

C:Accession: T09164

R:Simon, P.

submitted to the EMBL Data Library, January 1997

A:Reference number: 216599

A:Accession: T09164

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-323 <SIM>

A:Cross-references: EMBL:Y10465

A:Experimental source: subspecies Nobel

C:Genetics:

A:Note: prx14

C:Superfamily: peroxidase

C:Keywords: heme; iron; metalloprotein; oxidoreductase

F:37-118/Disulfide bonds: #status predicted

F:64/Active site: Arg #status predicted

F:68,196/Binding site: heme iron (His) (axial ligands) #status predicted

F:70-75/Disulfide bonds: #status predicted

F:124-319/Disulfide bonds: #status predicted

F:203-228/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 60.6%; Score 972.5; DB 2; Length 323;

Matches 197; Conservative 39; Mismatches 76; Indels 11; Gaps 4;

QY 1 MASPTLMQCLVAVSLSCVAHAQ---LSPTFYASCPNIOSTYRAMTQAVASEOMGMSLRPFHDFVGC 56

1 MASPTLMQCLVAVSLSCVAHAQ---LSPTFYASCPNIOSTYRAMTQAVASEOMGMSLRPFHDFVGC 60

57 SLRLEFPHDFVGCDSIILD---AGGKTAGPNLNSVGFVIDITIKRNVPAACPG 111

61 SLRLEFPHDFVGCDSIILD---AGGKTAGPNLNSVGFVIDITIKRNVPAACPG 120

112 VVSCADIALAARDGTNLGGFTWVPLGRDSTTASASLANSPPTASLGLTISLFG 171

121 TVSCADIALAARDGTNLGGFTWVPLGRDSTTASASLANSPPTASLGLTISLFG 180

172 RQGLSPDMTALSGAHTIGARCTFRGRIYGDPTDINAFALRQTCPRSGGCGNLAPIDVOPVPRFDYFTNL 229

181 RQGLSPDMTALSGAHTIGARCTFRGRIYGDPTDINAFALRQTCPRSGGCGNLAPIDVOPVPRFDYFTNL 240

230 PIDVOPVPRFDYFTNLNRGLFHSODELFNGSODALVROYTSASLSFNADFYAAMIRMGVGLTGTACQIRRN 289

241 PMDIOGTNTFENDYKYNLVAKRGLHSDQELFNGTQDALVRYTSNNAATFATDFAAAMVAMGNIISLTGTNEIRN 300

QY 290 RMGNVGLTGTACQIRRNCRVYN 312

Db 301 RMGDLKPLGTGTNEIRNCRVYN 323

## RESULT 4

OPNR7

peroxidase (EC 1.11.1.7) - turnip

N:Alternate names: hydrogen peroxide oxidoreductase; peroxidase 7

C:Species: Brassica rapa (turnip)

C:Date: 31-Oct-1980 #sequence\_revision 23-Oct-1981 #text\_change 03-Mar-2000

C:Accession: A91094; A91246; B23116; P06121; A00503

R:Mazza, G.; Wellinder, K.G.

Eur. J. Biochem. 108, 481-489, 1980

A:Title: Covalent structure of turnip peroxidase 7. Cyanogen bromide fragments, compl

A:Reference number: A91094; MUID:81003872; PMID:7408864

A:Accession: A91094

A:Molecule type: protein

A:Residues: 1-296 <MAZ>

A:Experimental source: cv. Blanc dur d'hiver

A:Note: the protein shown, TP7, is the principal isoperoxidase during winter in turni

R:Wellinder, K.G.; Mazza, G.

Eur. J. Biochem. 73, 353-358, 1977

A:Title: Amino-acid sequences of heme-linked, histidine-containing peptides of five p

A:Reference number: A91246; MUID:77138218; PMID:849740

A:Accession: A91246

A:Molecule type: protein

A:Residues: 32-65;161-175 <WEL>

A:Note: these two histidine-containing tryptic peptides are essential to the peroxida

R:Wellinder, K.G.

Eur. J. Biochem. 151, 497-504, 1985

A:Title: Plant peroxidases: their primary, secondary and tertiary structures, and rel

A:Reference number: A23116; MUID:85285060; PMID:2992968

A:Accession: B23116

A:Status: preliminary

A:Molecule type: protein

A:Residues: 7, 2-296 <WEL>



R:Diehn, S.H.; Burkhardt, W.; Graham, J.S.

Biochem. Biophys. Res. Commun. 195, 928-934, 1993

A:Title: Purification and partial amino acid sequence of wound-inducible, developmentally

A:Reference number: PM0612; M01D:93384622; PMID:8396932

A:Accession: PM0612

A:Molecule type: protein

A:Residues: 72, 2-32, 7AG, 35-88, 1D, 90-156, 1K, 158-296 <DIE>

C:Superfamily: peroxidase

C:Keywords: chromoprotein; glycoprotein; heme; iron; metalloprotein; oxidoreductase; py

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:11-91,44-49,97-292,176-201/Disulfide bonds: #status experimental

F:38/Active site: Arg #status predicted

F:42/Binding site: heme iron (His) (axial ligands) #status predicted

F:185/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 58.7%; Score 942; DB 1; Length 296;

Best Local Similarity 61.3%; Pred. No. 6, 2e-69;

Matches 182; Conservative 42; Mismatches 66; Indels 6; Gaps 2;

QY 23 QLSPTFVASCPCNLQSIYRAAMTOAVASEQRMGASLRLFFHDCPVGCGDSILLD---- 78

Db 1 QLTFTNYSTSCPNLSTYKSGVSAVSSQPRMGASTLRFHDCPVNCGDSILLDPTSS 60

QY 79 AGGKTAGPNLNSVGFVYDITKRNVEACPGVSCADITLALARDGTNLGGPTMSVP 138

Db 61 FTGEQVAGPNNRSARGFTVINDIKSAVERKCPGVSCADITLALARDSVVLGGSPNNVK 120

QY 139 LGRDSTVASASLANSPPEPTASTLGLTSLFGRQGLSPRDMTALSGAHTTGAQCTTFR 198

Db 121 VGRDQKFTASQAAANSNIPAPSMLSQLTSSFAVSLSTRDVALSGAHTTIGQRCVNR 180

QY 199 GRITGPTDINAFALRQCTPRSS--GGDGNLPIVQVPEFPAAYRTNLSRGLFHS 256

Db 181 ARVYENETNNAFALRQCTPRSS--GGDGNLPIVQVPEFPAAYRTNLSRGLFHS 240

QY 257 DDELNGSQDALVROYSASASLNFADVFAMIRMGVGLTGTAGQIRNCRVYN 312

Db 241 DQVLENGSGDSTIVRGYSNPSFSDFAAMIKMGDISPLTGGSGEIRKVCRTN 296

RESULT 5

T10252

peroxidase (EC 1.11.1.7) - radish

C:Species: Raphanus sativus (radish)

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000

C:Accession: T10252

R:Park, J.

Submitted to the EMBL Data Library, September 1995

A:Reference number: Z16998

A:Accession: T10252

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-315 <PAR>

A:Cross-references: EMBL:X91172

C:Experimental source: cultivar Handsome Hall

C:Genetics:

A:Gene: prx1

A:Insertion: 68/3; 131/3; 186/3

C:Superfamily: peroxidase

C:Keywords: chromoprotein; glycoprotein; heme; iron; oxidoreductase

Query Match 58.3%; Score 935; DB 2; Length 315;

Best Local Similarity 59.9%; Pred. No. 2, 5e-68;

Matches 184; Conservative 44; Mismatches 75; Indels 4; Gaps 2;

QY 10 LVAVSLSCVAHROLSPTFVASCPCNLQSIYRAAMTOAVASEQRMGASLRLFFHDCPV 69

Db 9 VIVITLLDGGPALTFTNYSTSCPNLSTYKSGVSAVSSQPRMGASTLRFHDCPVN 68

QY 70 GCDGSIILD--AGGKTAGPNLNSVGFVYDITKRNVEACPGVSCADITLALARDG 126

Db 69 GCDGSIILDDTSTFTGEQVAGPNNRSARGFTVINDIKSAVERKCPGVSCADITLALARDG 128

QY 127 TNLGGPTMSVPLGRDSTTASASLANSPPEPTASTLGLTSLFGRQGLSPRDMTALSGA 186

Db 129 VVGLGSPNNVNRARGDRAKTAASQAAANSNIPAPSMLSQLTSSFAVSLSTRDVALSGA 188

QY 187 HTTGACRTTFGRITNGRTDINAFALRQCTPRSG--GDGLAFIDVQTVREYTYFT 245

Db 189 HTTGACRTTFGRITNGRTDINAFALRQCTPRSG--GDGLAFIDVQTVREYTYFT 248

QY 246 NLSRGLFHSDELFNGSQDALVROYSASASLNFADVFAMIRMGVGLTGTAGQIR 305

Db 249 NLMAQGLHSDQVLFNGSGDSTIVRGYSNPSFSDFAAMIKMGDISPLTGGSGEIR 308

QY 306 RNCRVYN 312

Db 309 KVGGRTN 315

RESULT 6

S22087

peroxidase (EC 1.11.1.7) precursor - rice

C:Species: Oryza sativa (rice)

C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 04-Mar-2000

C:Accession: S22087

R:Reimann, C.; Ringli, C.; Dudler, R.

Submitted to the EMBL Data Library, June 1992

A:Description: cDNA cloning and sequence analysis of a pathogen-induced peroxidase fr

A:Reference number: S22087

A:Accession: S22087

A:Molecule type: mRNA

A:Residues: 1-317 <RE>

A:Cross-references: EMBL:X66125; M1D:g20285; P1D:CAA46916.1; P1D:g20286

C:Experimental source: cv. Nohrin

C:Superfamily: peroxidase

C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase

F:63/Active site: Arg #status predicted

F:67/Binding site: heme iron (His) (axial ligands) #status predicted

F:69-74/Disulfide bonds: #status predicted

F:117-312/Disulfide bonds: #status predicted

F:196-221/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 57.9%; Score 928; DB 2; Length 317;

Matches 186; Conservative 44; Mismatches 82; Indels 6; Gaps 3;

QY 1 MASPT---LMQVLAVSLSCVAHROLSPTFVASCPCNLQSIYRAAMTOAVASEQRMGAS 57

Db 1 MASATVSSLSLMLLVAAAMASVSAHOLSAFTFTDSCPNLSTYKSVITRAVASEARMGAS 60

QY 58 LRLFFHDCPVGCGDSILLDAGEKTAGPNLNSVGFVYDITKRNVEACPGVSCAD 117

Db 61 LRLFFHDCPVGCGDASVLL--SGDEQNAAGPVNGLSGFVINDAKARVEALICNPTVSCAD 119

QY 118 ILALARDGTNLGGPTMSVPLGRDSTTASASLANSPPEPTASTLGLTSLFGRQGLSP 177

Db 120 ILAVARDSVALGGPSMTVLGRRDSTTASALAMTDLPARSSSLAELINFGSKGDA 179

QY 178 RDMFTALSGAHTTIGQARTTFGRITNGRTDINAFALRQCTPR--SGDGLAFIDVQ 235

Db 180 TDVAVSLSGAHTTIGQACQCFRDRITNETNIDSAFAVQKQANCPRTGSDSLAVDTT 239

QY 236 PVRFDTAYFTNLSRGLFHSDELFNGSQDALVROYSASASLNFADVFAMIRMGVNG 295

Db 240 PNFADNAYYSNLSKGLHSDQVLFNGSGDSTIVRGYSNPSFSDFAAMIKMGDISPLTGGSGEIR 299

QY 296 VITGTAGQIRNCRVYN 313

Db 300 PLTGTGQIRNCRVYN 317

RESULT 7

T03329

peroxidase (EC 1.11.1.7) - rice



```

||||| 117 179
Db 120 DILVAARDVVALGPTWVLLGRDSTASTMAENDLPPTFDQNTTLTGKQLS
Oy 177 PRDMTALSGAHTTIGQACTFERGRIYGDTDINAFALROOTCPRSGDGNLAPIDYQRP 236
Db 180 MTBVALSGAHTTIGQSCRFRRIRYETININTFATSLRANCPGSGDSSLAPLDQTP 239
Oy 237 VRDPTAVFTNLSRRGLFHSDDQLFNGSGODALVROYASASLNFADPVAAIRMGVAV 296
Db 240 NGFDNAVYTNLMQKGLHSDQVLFNGGADNTRFSSSAAFTFAMVMGNIAIP 299
Oy 297 LTGTAGQIRRCRCRVNS 313
Db 300 KTGTGGQIRLVCSRVNS 316

```

## RESULT 10

```

S61408
peroxidase (EC 1.11.1.7) 4 precursor - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 27-Apr-1996 #sequence_revision 26-Jul-1996 #text_change 04-Mar-2000
C:Accession: S61408; S53110
R:Baga, M.; Chibbar, R.N.; Kartha, K.K.
Plant Mol. Biol. 29, 647-662, 1995
A:Title: Molecular cloning and expression analysis of peroxidase genes from wheat.
A:Reference number: S61405; MUID:96128008; PMID:8541492
A:Accession: S61408
A:Molecule type: DNA
A:Residues: 1-319 <BAG>
A:Cross-references: EMBL:X85230; NID:9732975; PIDN:CA59487.1; PID:9732976
A>Note: exon/intron junctions have been confirmed by mRNA sequencing
C:Genetics:
A:Insertions: 70/3; 134/3
C:Superfamily: peroxidase
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase; pyroglyutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-319/Product: peroxidase 4 #status predicted <MAG>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:34-114/Disulfide bonds: #status predicted
F:61/Active site: Arg #status predicted
F:65,192/Binding site: heme iron (His) (axial ligands) #status predicted
F:67-72/Disulfide bonds: #status predicted
F:120-315/Disulfide bonds: #status predicted
F:199-224/Disulfide bonds: #status predicted

```

```

Query Match 55.8%; Score 895.5; DB 2; Length 319;
Best Local Similarity 57.5%; Pred. No. 4e-65; Indels 9; Gaps 3;
Matches 181; Conservative 38; Mismatches 87; Indels 9; Gaps 3;

```

```

Oy 7 MOCIVAVSLSCVA--HMOISPTFYASSCPNLOSIVRAAMTQAAVSEORMGASLRLFF 63
Db 5 MASSISVLLILCLIAAPSSAQLSPRFARSCPRAQAIIRGVAAAVRSERMAASLRLHF 64
Oy 64 HDCFVCGDGSILDD---AGGEKTAGPNLNSVGRFEVIDTKRVEAACPGVSCADIL 119
Db 65 HDCFVCGDASILSDTATFTGEOGAGPNAGSIRGKNVIDNKAEVAVCTQTVSCADIL 124
Oy 120 ALAARCTNLGGPTWSVPLGRDSTTASASLANSPPTASISGLTLLSFGROGSPRD 179
Db 125 AAVARSVALGGPSWTVPLGRDSTTASASLANSPPTASISGLTLLSFGROGSPRD 184
Oy 180 MTAALSGAHTTIGQACTFERGRIYGDTDINAFALROOTCPR--SGGDGNLAPIDYQRP 237
Db 185 MTAALSGAHTTIGQACQNFRRIRYETININTFATSLRANCPRTGSGDSSLAPLDQTP 244
Oy 238 RPTAVFTNLSRRGLFHSDDQLFNGSGODALVROYASASLNFADPVAAIRMGVAV 297
Db 245 AFPNNAVYTNLMQKGLHSDQVLFNGGADNTRFSSSAAFTFAMVMGNIAIP 304
Oy 298 TGTAGQIRRCRCRVN 312
Db 305 TGTGGQIRLVCSRVN 319

```

## RESULT 11

```

B56555
peroxidase (EC 1.11.1.7), anionic, precursor - wood tobacco
C:Species: Nicotiana sylvestris (wood tobacco)
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 04-Mar-2000
C:Accession: B56555
R:Crivul, M.C.; Plesse, B.; Durr, A.; Marbach, J.; Parentier, Y.; Jamet, E.; Fleck,
Mech. Dev. 38, 121-132, 1992
A:Title: Characterization of genes expressed in mesophyll protoplasts of Nicotiana sy
A:Reference number: A56555; MUID:93041285; PMID:1419848
A:Accession: B56555
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-322 <CR1>
A:Cross-references: GB:M74103; NID:9170202
A:Experimental source: protoplasts
A>Note: sequence extracted from NCBI backbone (NCBIN:118061, NCBI:P:118062)
C:Superfamily: peroxidase
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-118/Disulfide bonds: #status predicted
F:63/Active site: Arg #status predicted
F:69,196/Binding site: heme iron (His) (axial ligands) #status predicted
F:71-76/Disulfide bonds: #status predicted
F:124-318/Disulfide bonds: #status predicted
F:203-229/Disulfide bonds: #status predicted

```

## Query Match

```

55.7%; Score 894; DB 2; Length 322;
Best Local Similarity 56.2%; Pred. No. 5.4e-65; Indels 10; Gaps 4;
Matches 181; Conservative 49; Mismatches 82; Indels 10; Gaps 4;

```

```

Oy 1 MASPT----LMOCIVAVSLSCVA--HMOISPTFYASSCPNLOSIVRAAMTQAAVSEORMG 55
Db 1 MNTPTQFAKAAAFSLLSLSCQCHQNSATFYDMTCPPALNTIRTSVQAISERRMA 60
Oy 56 ASLIRLFHDFVCGDGSILDD---AGEKTRAGPNLNSVGRFEVIDTKRVEAAACG 111
Db 61 ASLIRLFHDFVCGDGSILDDTETPSISEKALPLMGARFGIIEKKRVEKICP 120
Oy 112 VVSCADILVAARDVVALGPTWVLLGRDSTTASASLANSPPTASISGLTLLSFG 171
Db 121 VVSCADILVAARDVVALGPTWVLLGRDSTTASAKTLAEVDLPPTPLRLISSFA 180
Oy 172 ROGLSPRDMTALSGAHTTIGQACTFERGRIYGD--TDINAFALROOTCPRSGDGNLAP 230
Db 181 SKGLSTRDMVALSGAHTTIGQACFLFRIRYETININTFATSLRANCPRTGSGDSSLAP 240
Oy 231 IDVOTPVAFPTAVFTNLSRRGLFHSDDQLFNGSGODALVROYASASLNFADPVAAIR 290
Db 241 LDVTPPNQFDNNYRKNIQKKGLLSDQVLFNGGADNTRFSSSAAFTFAMVMGNIAIP 300
Oy 291 MGNVGLTGTAGQIRRCRCRVN 312
Db 301 MDLISPLSGNGITIRKVGGSVN 322

```

## RESULT 12

```

S61406
peroxidase (EC 1.11.1.7) 2 precursor - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 27-Apr-1996 #sequence_revision 26-Jul-1996 #text_change 04-Mar-2000
C:Accession: S61406; S53108
R:Baga, M.; Chibbar, R.N.; Kartha, K.K.
Plant Mol. Biol. 29, 647-662, 1995
A:Title: Molecular cloning and expression analysis of peroxidase genes from wheat.
A:Reference number: S61405; MUID:96128008; PMID:8541492
A:Accession: S61406
A:Molecule type: DNA
A:Residues: 1-314 <BAG>
A:Cross-references: EMBL:X85228; NID:9732971; PIDN:CA59485.1; PID:9732972
A>Note: exon/intron junctions have been confirmed by mRNA sequencing
C:Genetics:

```

A:introns: 70/3; 129/3  
 C:Superfamily: peroxidase  
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase; pyroglutamic acid  
 F:1-23/Domain: signal sequence #status predicted <Sig>  
 F:24-314/Product: peroxidase 2 #status predicted <Sig>  
 F:24-314/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
 F:34-109/Disulfide bonds: #status predicted  
 F:61/Active site: Arg #status predicted  
 F:65-187/Binding site: heme iron (His) (axial ligands) #status predicted  
 F:115-309/Disulfide bonds: #status predicted  
 F:194-220/Disulfide bonds: #status predicted

Query Match 54.1%; Score 867; DB 2; Length 314;  
 Best local Similarity 56.4%; Pred. No. 8e-63;

Matches 172; Conservative 50; Mismatches 81; Indels 2; Gaps 2;

QY 10 LVAVSLSCVAHAQSPTFYASCPNLQSIIVAAAMTOAVASQKAGASLLRFHDCFPVQ 69  
 11 LVAVSLSCVAHAQSPTFYASCPNLQSIIVAAAMTOAVASQKAGASLLRFHDCFPVQ 70  
 QY 70 GCDGSLDAGEKTAAGPMLNSVGFVDTIKRNVAAACPGVSCADITLAAADGTNL 129  
 71 GCDASVLL-SGMQNAAGPVGSLRFGVYDNTIKTQLESCKQTVSCADITLVAARDVVA 129  
 QY 130 LGPPTMSVPLGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSPRDMTALSGARTI 189  
 130 LGPPTMSVPLGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSPRDMTALSGARTI 189  
 QY 190 GOARCTFGRITV-GPDTINASFALRQOTCPRSGDGNLAPIDVQTVPRDTATFTL 248  
 190 GAAQCSNFRITRYGCDTNTAFATSLKANCPOSGGNLLENLDTTNAFDMAYTNL 249  
 QY 249 SRRLGHSQELFNGSODALVROYASASLFNADFAAMIRMGVNLGTAGQIRNRC 308  
 250 SOKGLHSQVLEFNNDTDTNFRNFASNAAFSSAFTTAMIKMGIAPLTGQIRLSC 309  
 QY 309 RYVNS 313  
 310 SKVNS 314

## RESULT 13

S13325

peroxidase (EC 1.11.1.7) precursor - wheat

C:Species: Triticum aestivum (common wheat)

C&gt;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 04-Mar-2000

C:Accession: S13325

R:Hertig, C.; Rebmann, G.; Bull, J.; Mauch, F.; Dudler, R.

A:Title: Sequence and tissue-specific expression of a putative peroxidase gene from wheat

A:Reference number: S13325; MUID:9136838; PMID:1653627

A:Accession: S13325

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-312 &lt;HER&gt;

A:Cross-references: GB:X53675; GB:S54871; NID:921830; PIDN:CA37713.1; PID:921831

C:Genetics:

A:Introns: 68/3; 127/3

C:Superfamily: peroxidase

C:Keywords: chromoprotein; glycoprotein; heme; iron; metalloprotein; oxidoreductase

F:1-23/Domain: signal sequence #status predicted &lt;Sig&gt;

F:24-312/Product: peroxidase #status predicted &lt;Sig&gt;

F:34-107/Disulfide bonds: #status predicted

F:61/Active site: Arg #status predicted

F:65-185/Binding site: heme iron (His) (axial ligands) #status predicted

F:113-307/Disulfide bonds: #status predicted

F:192-218/Disulfide bonds: #status predicted

Query Match

Best local Similarity 53.0%; Score 849.5; DB 2; Length 312;

Matches 173; Conservative 51; Mismatches 85; Indels 5; Gaps 4;

QY 1 MASPTLMQCLVAVSLSCVAHAQSPTFYASCPNLQSIIVAAAMTOAVASQKAGASLLRF 60  
 3 MGSASISLVLVA-LATVAAGQSLSPTFYDTSCRALVAIKSGVAAVSDPRMGASLLRFHDCFPVQ 61  
 QY 61 LFHDCFPVQGDGSLDAGEKTAAGPMLNSVGFVDTIKRNVAAACPGVSCADITL 120  
 62 LFHDCFPVQGDGSLDAGEKTAAGPMLNSVGFVDTIKRNVAAACPGVSCADITL 118  
 QY 121 LAARDGTLNLCGPTMSVPLGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSPRDM 180  
 121 LAARDGTLNLCGPTMSVPLGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSPRDM 180  
 QY 119 VAARDVVALCGPSMTVPLGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSPRDM 178  
 119 VAARDVVALCGPSMTVPLGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSPRDM 178  
 QY 181 TALSGARTIIGARCTTFRGRTV-GPDTINASFALRQOTCPRSGDGNLAPIDVQTVPR 239  
 179 VALSGARTIIGARCTTFRGRTV-GPDTINASFALRQOTCPRSGDGNLAPIDVQTVPR 238  
 QY 240 DTAVFTNLISRRGLFHSQDELFGSODALVROYASASLFNADFAAMIRMGVNLGTAGQIRNRC 239  
 239 DTAFTNLISRRGLFHSQDELFGSODALVROYASASLFNADFAAMIRMGVNLGTAGQIRNRC 238  
 QY 300 TAGQIRNRCRYVNS 313  
 299 TAGQIRNRCRYVNS 312

## RESULT 14

T09165

probable peroxidase (EC 1.11.1.7) (clone pc18) - spinach (fragment)

C:Species: Spinacia oleracea (spinach)

C&gt;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 20-Jun-2000

C:Accession: T09165

R:Simon, P.

A:Reference number: T09165

A:Accession: T09165

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-315 &lt;SIM&gt;

A:Cross-references: EMBL:Y10466

A:Experimental source: subspecies Nobel

C:Genetics:

A:Gene: prx15

C:Superfamily: peroxidase

C:Keywords: heme; iron; metalloprotein; oxidoreductase

F:1-111/Disulfide bonds: #status predicted

F:58/Active site: Arg #status predicted

F:62-189/Binding site: heme iron (His) (axial ligands) #status predicted

F:117-311/Disulfide bonds: #status predicted

F:196-221/Disulfide bonds: #status predicted

Query Match

Best local Similarity 52.8%; Score 846.5; DB 2; Length 315;

Matches 170; Conservative 42; Mismatches 85; Indels 5; Gaps 2;

QY 16 LSCVAHAQSPTFYASCPNLQSIIVAAAMTOAVASQKAGASLLRFHDCFPVQGDGSI 75  
 14 LSLTVNGQLSPNFSYSCPCNALRIKQGIKRIKKEARVAGASLLRFHDCFPVQGDGSI 73  
 QY 76 LLD-AGAGTAPGNLNSVGFVDTIKRNVAAACPGVSCADITLAAADGTNL 131  
 74 LLDSTFGEKTAAGPMLNSVGFVDTIKRNVAAACPGVSCADITLAAADGTNL 133  
 QY 132 GPTMSVPLGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSPRDMTALSGARTI 191  
 134 GPTMSVPLGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSPRDMTALSGARTI 193  
 QY 192 ARCTTFRGRTV-GPDTINASFALRQOTCPRSGDGNLAPIDVQTVPRDTATFTL 251  
 194 ARCTTFRGRTV-GPDTINASFALRQOTCPRSGDGNLAPIDVQTVPRDTATFTL 253  
 QY 252 GLFHSQDELFGSODALVROYASASLFNADFAAMIRMGVNLGTAGQIRNRCRV 310

```

Db      254  GLLHSQQLXSGNNADAYVKKYASKSGEFPQFEGNSMTIRMGNIKPLTIGTGQIRRCRK 3133
          || |||||:::| : || |||:::| : : : |||||::| ||||| |||||
Qy      311  VN 312
          |
Db      314  SN 315

```

## RESULT 15

peroxidase [EC:1.11.1.7] precursor, pathogen-induced - wheat  
C:Species: Triticum aestivum (common wheat)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 04-Mar-2000  
C:Accession: S13375  
R:Rehmann, G.; Hertig, C.; Bull, J.; Mauch, F.; Dudler, R.  
Plant Mol. Biol. 16, 329-331, 1991  
A>Title: Cloning and sequencing of cDNAs encoding a pathogen-induced putative peroxidase  
A:Reference number: S13375; M01D:91370874; PMID:1893103  
A:Accession: S13375  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-312 <NB>  
A:Cross-references: GB:X56011; GB:S55687; NID:q21828; PIDD:CAA9486.1; PTD:q21829  
A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 135-Tyrosine  
C:Superfamily: peroxidase  
C:Keywords: chromoprotein; glycoprotein; heme; iron; metalloprotein; oxidoreductase  
F:1-32/Domain: signal sequence #status predicted <Sig>  
F:32-312/Product: peroxidase #status predicted <Mat>  
F:33-108/Disulfide bonds: #status predicted  
F:60/Active site: Arg #status predicted  
F:64,186/Binding site: heme iron (His) (axial ligands) #status predicted  
F:66-71/Disulfide bonds: #status predicted  
F:114-307/Disulfide bonds: #status predicted  
F:193-218/Disulfide bonds: #status predicted

Query Match	52.3%;	Score 838.5;	DB 2;	Length 312;
Best Local Similarity	54.0%;	Pred. No. 1.6e-60;		
Matches .169;	Conservative 49;	Mismatches 94;	Indels 1;	Gaps 1;

QY	1	MASPTLMOCIVAVSLLSCVAAHOLSPFFVYASCCPNCLOSTIYRAMOAVQASEDRGASILR	60
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	1	MAASACSLSYLVVALATFASASQSLPFFIYDPSCRALATIKKSVAAVANSDDPRMCASILR	60
QY	61	LFFHDCFCVOCGDSILLDAGCKTAPNLNSVRCGEYIDTIRKNVEACPGVWSCADILIA	120
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	61	LHFHDFCFVOCGCAAVLL - SGEMONALINAGSLRGFGVIDSIKFOIETALCONQFSCADILIT	119
QY	121	LAARDGTNLLGGVYNSVPLGRDRDSTFASASLANSNPPPPASLGTILISLFGOGSLSPEDM	180
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	120	VAARDVSVALGGSWSVPLGRDRDSTIDANEAANSDDLPGLTSSASDDELAFRRKGLITLIDM	179
QY	181	TALSGAHTTIGQARCTTFRCGRITGDTIDINASFALROOTCFRSGGDNLAIPIDVOTPVRED	240
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	180	VALSGAHTTIGQAOCCGFKRITRYNETNIDTVAFTSLRANCPSRMSGDSLAMIDTTAANTFD	239
QY	241	TAFVNTLNSRGLFHHDELFNGSGODALVROYSSASLENADVFVAAMIIRMGVNVGLTGR	300
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	240	NAYVTNLMSQKGLSHSDQYLEFNNDITDVTANFASNAAPSSFTTAMIKGNLAPKGT	299
QY	301	AGQIRNRCRYVNS	313
		: : : : :    : : : : :    : : : : :    : : : : :    : : : : :	
Db	300	OGQIRLSCRYVNS	312

Search completed: April 26, 2003, 12:35:37  
Job time : 47 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:22:23 ; Search time 25 Seconds

(without alignments)  
519.284 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 1604  
Sequence: 1 MASPTLMQCLVAVSLSCVA.....VGVLTGTAGQIRNCRVNS 313

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	942	58.7	296	1	PERX_BRARA
2	927	57.8	316	1	PERL_ARATY
3	926	57.7	314	1	P37835 oryza sativ
4	894	55.7	322	1	PERX_NICSY
5	849.5	53.0	312	1	PERL_WHEAT
6	818	51.0	315	1	PERL_HORVU
7	736.5	45.9	305	1	PERX_ARMRU
8	723	45.1	324	1	PERX_TOBAC
9	713.5	44.5	332	1	PERC_ARMRU
10	707	44.1	353	1	PERA_ARMRU
11	702.5	43.8	351	1	PERB_ARMRU
12	690	43.0	354	1	PERC_ARATH
13	684	42.6	349	1	PERE_ARATH
14	679	42.3	349	1	PERX_ARMRU
15	676.5	42.2	347	1	PER2_ARMRU
16	640	39.9	326	1	PERL_ORYSA
17	638	39.8	330	1	PER2_ARATY
18	600	37.4	292	1	PER2_CUCSA
19	550.5	34.3	363	1	PER2_LYCPS
20	546.5	34.1	364	1	PER1_LYCPS
21	535.5	33.4	351	1	PERX_SOLFU
22	454.5	28.3	158	1	PERX_LUPPO
23	262	16.3	170	1	PER2_HORVU
24	206	12.8	80	1	PERX_WHEAT
25	197.5	12.3	249	1	APX1_PEA
26	190.5	11.9	249	1	APX1_PEA
27	129	8.0	382	1	PER4_PHACH
28	118.5	7.4	378	1	PER1_PHACH
29	112	7.0	361	1	LIG_PHURA
30	108.5	6.8	364	1	PER_ARTRA
31	107.5	6.7	366	1	PER_CORCT
32	106.5	6.6	366	1	TRUB_TREPA
33	104.5	6.5	372	1	LIG4_PHACH

34	101	6.3	372	1	LIGC_TRAVE	P20013 trameles ve
35	99	6.2	486	1	CUG1_HUMAN	Q92879 homo sapien
36	99	6.2	486	1	CUG1_MOUSE	P28659 mus musculu
37	97	6.0	517	1	HEX_ADE08	P36852 human adeno
38	96	6.0	730	1	CAT4_HALMA	O59651 haloarcula
39	95	5.9	372	1	LIG6_PHACH	P50622 phanerocha
40	95	5.9	397	1	TYRB_ECOLI	P04693 escherichia
41	93	5.8	545	1	THSA_ARCFU	O28045 archaeoglob
42	91	5.7	1672	1	PMPB_CHLMU	O9pjy2 chlamydia m
43	90.5	5.6	349	1	TL29_ARATH	P82281 arabidopsis
44	90.5	5.6	1196	1	ICEV_PSEEX	O34479 pseudomonas
45	90	5.6	554	1	ILVD_THEMA	Q9wz21 thermocoga

## ALIGNMENTS

RESULT 1	PERX_BRARA	STANDARD:	PRT:	296 AA.
AC	P00434:			
DT	21-JUL-1986 (rel. 01, Created)			
DT	01-FEB-1994 (rel. 28, Last sequence update)			
DT	15-JUN-2002 (rel. 41, Last annotation update)			
DE	Peroxidase p7 (EC 1.11.1.7).			
OS	Brassica rapa (Turnip).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Nosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Brassica.			
OX	NCBI_TaxID=51350;			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN=cv. Blanc dur d'hiver;			
RX	MEDLINE=81003872; PubMed=7408864;			
RA	Mazza G., Welinder K.G.;			
RT	"Covalent structure of turnip peroxidase 7. Cyanogen bromide			
RT	fragments, complete structure and comparison to horseradish			
RT	peroxidase C.";			
RL	Eur. J. Biochem. 108:481-489(1980).			
RN	[2]			
RP	SEQUENCE OF 32-65 AND 161-175.			
RX	MEDLINE=7138218; PubMed=849740;			
RA	Welinder K.G., Mazza G.;			
RT	"Amino-acid sequences of heme-linked, histidine-containing peptides			
RT	of five peroxidases from horseradish and turnip.";			
RL	Eur. J. Biochem. 73:353-358(1977).			
CC	-1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.			
CC	-1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.			
CC	-1- MISCELLANEOUS: THE PROTEIN SHOWN, TP7, IS THE PRINCIPAL			
CC	ISOFORM OF THE PEROXIDASE FAMILY. PLANT PEROXIDASE			
CC	SUBFAMILY.			
CC	PIR: A00503; OPMN7.			
DR	PIR: B2316; B2316.			
DR	HSSP: P22195; 1SCH.			
DR	InterPro: IPR002016; Peroxidase.			
DR	Pfam: PF00141; Peroxidase; 1.			
DR	PRINTS: PR00458; PEROXIDASE.			
DR	PROSITE: PS00435; PEROXIDASE_1; 1.			
DR	PROSITE: PS00436; PEROXIDASE_2; 1.			
KW	Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;			
KW	Multigene family.			
FT	MOD_RES	1		
FT	ACT_SITE	38		
FT	ACT_SITE	42		
FT	ACT_SITE	169		
FT	METAL	169		
FT	DISULFID	11		
FT	DISULFID	91		
FT	DISULFID	44		
FT	DISULFID	49		
FT	DISULFID	97		
FT	DISULFID	176		
FT	CARBOHYD	185		
FT	SEQUENCE	296 AA;		
FT	SEQUENCE	31086 MW;		
FT	SEQUENCE	53C9CCE5B2A7937 CRC64;		

Query Match 58.7%; Score 942; DB 1; Length 296;  
 Best Local Similarity 61.5%; Pred. No. 3,8e-68;  
 Matches 182; Conservative 42; Mismatches 66; Indels 6; Gaps 2;

QY 23 QLSPTFAVSSCPNOSTIVRAMQAVASEORMGASLRLFFHDFVGGCDSTLDDTSS 78  
 DB 1 QLTNFTSTSCPNTLSTVSKSVASVSSQPMASIRLFFHDFVGGCDSTLDDTSS 60  
 QY 79 AGGKTAGPNTNSVGEVIDITIKRNYEACPGVASCADILALARDGTLNLSGGPTWSP 138  
 DB 61 FTGEGMGNPNNSARGFVINDIKSAVEKACPGVASCADILALARDGTLNLSGGPTWSP 120  
 QY 139 LGRDSTASLANSNPPTTASLGLTISFGROGSPRMTLSGAHTTIGQARCTTFR 198  
 DB 121 VGRADAKTASQAANSNPAPSMLSLQITSEFSAVGSTRDWVALSGHHTIGOSRCVNR 180  
 QY 199 GRVGDPTDINSFALRQOTCPRS--GGDGNLAPIDVQTPVRFPTAVFTNLSRGLFHS 256  
 DB 181 ARVNETINNAFATLRQSCPRAGSGDANLAPLINSATSPNSYFKNLMAORGLHS 240  
 QY 257 DOELFNGSGDALTVOYSASASLENAFVAAIMKGNVGTITGAGTIRRCRYVN 312  
 DB 241 DOVLENGSGSTDSIVRGYSNSESSENFDAAMIKMGDISPLTGSGELIKVCGKTN 296

RESULT 2  
 PER1 ARAHY STANDARD; PRT; 316 AA.  
 AC P22195;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cationic peroxidase 1 precursor (EC 1.11.1.7).  
 GN PNC1.  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Arachis.  
 OX NCBI\_Taxid=3818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-91062381; PubMed-2247460;  
 RA Buffard D., Breda C., van Huysstee R.B., Asemota O., Pierre M.,  
 RA Dang Ha D.B., Esnault R.;  
 RT "Molecular cloning of complementary DNAs encoding two cationic  
 RT peroxidases from cultivated peanut cells.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 87:8874-8878(1990).  
 RL [2]  
 RP REVISION TO 47.  
 RA Esnault R.;  
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RA MEDLINE-96398617; PubMed-8805359;  
 RA Schuller D.J., Ban N., van Huysstee R.B., McPherson A., Poulos T.L.;  
 RT "The crystal structure of peanut peroxidase.";  
 RT Structure 4:311-321(1996).  
 CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,  
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD  
 CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE  
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.  
 CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
 CC -1- COFACTOR: Binds 1 prothème IX, 1 iron(III) ion and 2 calcium ions  
 CC per subunit.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.1db-sib.ch/announce/>  
 or send an email to [license@sib-sb.ch](mailto:license@sib-sb.ch)).

CC EMBL: M37636; AAB06183.1; -  
 CC DR PIR: A38265; A38265.  
 CC DR PDB: 1SCH; 11-JUL-96.  
 CC DR GlycosultedB: P22195; -  
 CC DR InterPro: IPR002016; Peroxidase.  
 CC DR Pfam: PF00141; peroxidase\_1.  
 CC DR PRINTS: PR00458; PEROXIDASE\_1.  
 CC DR PROSITE: PS00435; PEROXIDASE\_1.  
 CC DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 CC KM Oxidoreductase; Glycoprotein; peroxidase; Iron; Heme; Calcium-binding;  
 CC KM Multigene family; Signal; 3D-structure.  
 CC FT SIGNAL 1 22  
 CC FT CHAIN 23 316  
 CC FT CA\_BIND 65 65  
 CC FT CA\_BIND 68 68  
 CC FT CA\_BIND 70 70  
 CC FT CA\_BIND 72 72  
 CC FT CA\_BIND 74 74  
 CC FT CA\_BIND 86 86  
 CC FT CA\_BIND 192 192  
 CC FT CA\_BIND 236 236  
 CC FT CA\_BIND 239 239  
 CC FT CA\_BIND 242 242  
 CC FT CA\_BIND 244 244  
 CC FT ACT SITE 60 60  
 CC FT ACT SITE 64 64  
 CC FT METAL 191 191  
 CC FT MOD\_RES 23 23  
 CC FT CARBOHYD 82 82  
 CC FT CARBOHYD 166 166  
 CC FT DISULFID 33 113  
 CC FT DISULFID 66 71  
 CC FT DISULFID 119 312  
 CC FT DISULFID 198 223  
 CC SQ SEQUENCE 316 AA; 33517 MW; 2CC271F8EBB8C9F0 CRC64;

Query Match 57.8%; Score 927; DB 1; Length 316;  
 Best Local Similarity 58.5%; Pred. No. 6,4e-67;  
 Matches 182; Conservative 41; Mismatches 78; Indels 10; Gaps 2;

QY 6 LMCLVAVSLSCVAHQSLPTFASSCPNOSTIVRAMQAVASEORMGASLRLFFHD 65  
 DB 12 IFMCLIG-----LGSALSSNFEYATKCPNALSTKSAVSAVAKEMGASLRLFFHD 65  
 QY 66 CFVGGCDSTLDDTSSNFTGKTNAGTPNANSIRGEVIDITIKSYESLCPGVSCADILAV 125  
 DB 66 CFVGGCDASVLLDTSNFTGKTNAGTPNANSIRGEVIDITIKSYESLCPGVSCADILAV 125  
 QY 122 AARDGTLNLSGGPTWSPVPLGRDSTTASLANSNPPTTASLGLTISLFGROGSPRDMT 181  
 DB 126 AARDSVVALGASVNNVLLGRDSTTASLANSNPPTTASLGLTISLFGROGSPRDMT 185  
 QY 182 ALSGAHTTIGQARCTTFRIGRTYGDIDINAFALRQOTCPSSGGGPNLAPIDVQTPVRFDT 241  
 DB 186 TLSGAHTTIGQACPAFRIRYNESNIDPTAKSLQANCPVSGGPTNLSPTVTTPNKFEN 245  
 QY 242 AFTNLSRRLFLSHDDELFNNGSODALVROYASASLENAFVAAIMKGNVGTITGTA 301  
 DB 246 AYTILNRKKKLLSHDQLEFNNGSODALVROYASASLENAFVAAIMKGNVGTITGTA 305  
 QY 302 GOIRRNCRVYN 312  
 DB 306 GOIRTNCRKTN 316

RESULT 3  
 PER2 ORYSA STANDARD; PRT; 314 AA.  
 AC P37835;



OY		240	DTAATTTNLSRGFHSQOEALVROXSASLEFNADVMANGVNGT	299
DB		242	DSAYITTNLMSKGLHSDQLVFNGSGTDNTVRNFSSNTAENSAFTAAVMKNISPLTG	301
OY		300	TAGOTIRNCRCRYVN	312
DB		302	TGGGRLCSKVYN	314
RESULT 4				
ID	PBRX_NICSY	STANDARD;	PRT;	322 AA.
AC	PERX_NICSY	002200;		
DT	01-JUL-1993	(Rel. 26, Created)		
DT	01-JUL-1993	(Rel. 26, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Lignin forming anionic peroxidase precursor (EC 1.11.1.7).			
OS	Nicotiana sylvestris (Wood tobacco).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotsyledons; core eudicots; Asterales; Asteraceae; Solanales; Solanaceae; Nicotiana. [1]			
RN	NCBI_TaxId=4096;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv_XANTHI NC;			
RX	MEDLINE=93041285; PubMed=1419848;			
RA	Criqui M.-C., Plesse B., Durr A., Warbach J., Parmentier Y., Jamet E., Fleck J.;			
RT	"Characterization of genes expressed in mesophyll protoplasts of Nicotiana sylvestris before the re-initiation of the DNA replicational activity.";			
RL	Mech. Dev. 38:121-132(1992).			
CC	-I FUNCTION: THIS ENZYME PLAYS AN INTEGRAL ROLE IN SECONDARY CELL WALL BIOSYNTHESIS BY THE POLYMERIZATION OF CINNAMYL ALCOHOLS INTO LIGNIN AND BY FORMING RIGID CROSS-LINKS BETWEEN CELULOSE, PECTIN, HYDROXY-POLINE-RICH GLYCOPROTINS, AND LIGNIN.			
CC	-I CATALYTIC ACTIVITY: Donor + H(2)O(2) -> oxidized donor + 2 H(2)O.			
CC	-I COFACTOR: Binds 1 protoheme IX and 1 Iron(iii) ion.			
CC	-I TISSUE SPECIFICITY: MESOPHYLL PROTOPLASTS AND TO A MUCH LESSER EXTENT, ROOTS AND GERMINATING SEEDS.			
CC	-I DEVELOPMENTAL STAGE: BEFORE RE-INITIATION OF THE DNA REPLICATIONAL ACTIVITY.			
CC	-I SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE SUBFAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/terms/ or send an email to license@ebi.ac.uk).			
DR	EMBL; M74103; AAA34050.1; .-			
DR	HSPB; P22195; ISCH.			
DR	InferPro: IPR002016; Peroxidase.			
DR	pfam; PF00141; peroxidase_1.			
DR	PRINTS; PR00458; PEROXIDASE.			
DR	PROSITE; PS00435; PEROXIDASE_1; 1.			
DR	PROSITE; PS00436; PEROXIDASE_2; 1.			
KW	Oxido-reductase; Peroxidase; Iron; Heme; Glycoprotein; Signal.			
FT	SIGNAL	1..27		
FT	CHAIN	28..322		
FT	MOD_RES	28..28		
FT	ACT_SITE	65..65	LIGNIN FORMING ANIONIC PEROXIDASE.	
FT	ACT_SITE	69..69	PYRROLIDONE CARBOXYLIC ACID	
FT	METAL	196..196	(BY SIMILARITY).  BY SIMILARITY.  DIGITAL HISTIDIENE (BY SIMILARITY).  IRON (PROTOHEME IX AXIAL LIGAND) (BY SIMILARITY).  BY SIMILARITY.  BY SIMILARITY.	
FT	DISULFID	38..118		
FT	DISULFID	71..76		
FT	DISULFID	124..318		



CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to a license@sib-sib.ch).  
 CC -----  
 DR EMBL: X56396; CAA41294.1; -  
 DR HSSP: P22195; 1SCH.  
 DR InterPro: IPR002016; Peroxidase.  
 DR Pfam: PF00141; peroxidase; 1.  
 DR PRINTS: PS00458; PEROXIDASE.  
 DR PROSITE: PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 KM Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;  
 KM Multigene family; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 1 21  
 FT ACT\_SITE 59 59  
 FT ACT\_SITE 63 63  
 FT METAL 186 186  
 FT DISULFID 32 107  
 FT DISULFID 65 70  
 FT DISULFID 113 310  
 FT DISULFID 193 219  
 FT CARBOHYD 158 158  
 FT CARBOHYD 265 265  
 SQ SEQUENCE 315 AA; 32976 MW; 4867179BD0296172 CRC64;  
 Query Match 51.0%; Score 818; DB 1; Length 315;  
 Best Local Similarity 53.6%; Pred. No. 3; 1e-58;  
 Matches 170; Conservative 54; Mismatches 87; Indels 6; Gaps 5;  
 QY 1 MASTLMOCCLVAVSLSCVHAQISPTFYASCPNLSQIYRAAMTOAVASEQRMGASLIR 60  
 DB 1 MASSSYSLVLAVALVT-AASAOISPTFYDTSCPRALATIKSGMAVTSDDPRGASLIR 59  
 QY 61 LFFHDCFYOGCDGSIILDADGEGKTAGPNLNSVRGEVIDTIKRNVENACGVVSCADILA 120  
 DB 60 LHFHDCFYOGCDASVLL-SCMEONAIIPNAGSLKRGFYIDSIKIOIEIKQYTSQADILR 118  
 QY 121 LAARDGNTLLGCPWVSVPLGRDSTTASASLANSNPPPTASLIGTLISLFGRO-GLSPRD 179  
 DB 119 VAARDSVVALGSPWVPLGRDSDIDANENDEANTDLPGFNSRAELFAFLKKGGLNTVD 178  
 QY 180 MTAISGAHTTGOARCTFRGRIT-GRNDINAFALAQOCTFRS--GGDGNLAPIDVQTP 236  
 DB 179 MVAISGAHTTGOACSTFRARITGGDTNINAAVASLARCPQVGVSGDSLANLDTTAA 238  
 QY 237 VRFDATFTNLISRRGLFHSODELFNGSODALVROYASASLNFADVFAMIRMGVNVG 296  
 DB 239 NTFPNAYTNTLMSQKGLHSDQVLFNNDTIDNTRNFASNPAAASSSETTAMIKMGVIA 298  
 QY 297 LTGTAGQIRNCRVYNS 313  
 DB 299 KTGTGQIRLSCSRVNS 315  
 RESULT 7  
 PERX\_ARMRU STANDARD; PRT; 305 AA.  
 AC P80679;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Peroxidase (EC 1.11.1.7).  
 OS Armoracia rusticana (Horseradish) (Armoracia laphatifolia).

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Armoracia.  
 CC NCBI\_TaxID=3704;  
 RN [1]  
 RP SEQUENCE.  
 RA Rasmussen C.B., Stoffer B., Wellinder K.G.;  
 RU Submitted (Aug-1996) to the SWISS-PROT data bank.  
 CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,  
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD  
 CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE  
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.  
 CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
 CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.  
 DR HSSP: P00433; 2ATJ.  
 DR InterPro: IPR002016; Peroxidase.  
 DR Pfam: PF00141; peroxidase; 1.  
 DR PRINTS: PS00458; PEROXIDASE.  
 DR PROSITE: PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 KM Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;  
 FT ACT\_SITE 42 42  
 FT ACT\_SITE 64 64  
 FT METAL 169 169  
 FT DISULFID 11 91  
 FT DISULFID 44 49  
 FT DISULFID 97 299  
 FT DISULFID 176 208  
 FT MOD\_RES 1 1  
 FT CARBOHYD 3 3  
 FT CARBOHYD 13 13  
 FT CARBOHYD 147 147  
 FT CARBOHYD 185 185  
 FT CARBOHYD 197 197  
 FT CARBOHYD 211 211  
 FT CARBOHYD 267 267  
 SQ SEQUENCE 305 AA; 31899 MW; 6A70039215C1737 CRC64;  
 Query Match 45.9%; Score 736.5; DB 1; Length 305;  
 Best Local Similarity 48.5%; Pred. No. 9; 6e-52;  
 Matches 147; Conservative 48; Mismatches 95; Indels 13; Gaps 3;  
 QY 23 QLSPTFYASCPNLSQIYRAAMTOAVASEQRMGASLIRLFFHDCFYOGCDGSIILDAGG- 81  
 DB 1 QLNATYSGTGPASATVASTIQOAFQSDTRIGASLIRLHFHDCFYOGCDASILLDDSGS 60  
 QY 82 ---EKTAGPNLNSVRGEVIDTIKRNVENACGVVSCADILALAARDGNTLLGPTWSP 138  
 DB 61 IQSEKNAGNANSARGFENVNDIKTALENTCGVYSCDILALASEASVSLGSPWVLL 120  
 QY 139 LGRDSTTASASLANSNPPPTASLIGTLISLFGROGLSRDMTALSAGHTTGOARCTFR 198  
 DB 121 LGRDSDITLANGLANSALSPREGSLNITKFSAYGLNNDNLVALISGAHTFRARCGVFN 180  
 QY 199 GRITG-----DTDNASFAALROOCTPRSGDGNLADIVQTPVRFDATFTNLISRR 251  
 DB 181 NRLFNSGTNGPPTLNTSLLSLQOLCPONGSASITTLMDLSTDPAPNNNFALNCRN 240  
 QY 252 GLFHSODELFN--GGSODALVROYASASLNFADVFAMIRMGVNVLTGTAGQIRNCR 309  
 DB 241 GLQSDQELFSTLGSATIAVVTISFAISNQTLEFOAFQASIMNMGNISPLTSGNGETRLDCK 300  
 QY 310 VVN 312  
 DB 301 KVD 303  
 RESULT 8  
 PERX\_TOBAC STANDARD; PRT; 324 AA.  
 ID PERX\_TOBAC

AC P11965;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lignin forming anionic peroxidase precursor (EC 1.11.1.7).  
 OS Nicotiana glauca (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lagrimini L.M., Burkhardt W., Moyer M., Rothstein S.;  
 RT Molecular cloning of complementary DNA encoding the lignin-forming  
 RT peroxidase from tobacco: molecular analysis and tissue-specific  
 RT expression.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7542-7546(1987).  
 CC -1- FUNCTION: THIS ENZYME PLAYS AN INTEGRAL ROLE IN SECONDARY CELL  
 CC WALL BIOSYNTHESIS BY THE POLYMERIZATION OF CINNAMYL ALCOHOLS  
 CC INTO LIGNIN AND BY FORMING RIGID CROSS-LINKS BETWEEN CELLULOSE,  
 CC PECTIN, HYDROXY-PROLINE-RICH GLYCOPROTEINS, AND LIGNIN.  
 CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
 CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J02979; AAA34108.1; -  
 DR PIR: A39889; A39889.  
 DR HSSP: P00433; 2ATY.  
 DR InterPro: IPR002016; Peroxidase.  
 DR Pfam: PF00141; peroxidase; 1.  
 DR PRINTS: PS00458; PEROXIDASE.  
 DR PROSITE: PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 KM Oxidoreductase; Peroxidase; Iron; Heme; Glycoprotein; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 324 LIGNIN FORMING ANIONIC PEROXIDASE.  
 FT MOD\_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.  
 FT ACT\_SITE 60 60 BY SIMILARITY.  
 FT ACT\_SITE 64 64 DISTAL HISTIDINE (BY SIMILARITY).  
 FT METAL 189 189 IRON (PROTOHEME IX AXIAL LIGAND) (BY  
 FT SIMILARITY).  
 FT DISULFID 33 111 BY SIMILARITY.  
 FT DISULFID 66 71 BY SIMILARITY.  
 FT DISULFID 117 320 BY SIMILARITY.  
 FT DISULFID 196 228 BY SIMILARITY.  
 FT CARBOHYD 35 35 N-LINKED (GLCNAc . . .) (POTENTIAL).  
 FT CARBOHYD 150 150 N-LINKED (GLCNAc . . .) (POTENTIAL).  
 FT CARBOHYD 207 207 N-LINKED (GLCNAc . . .) (POTENTIAL).  
 SQ SEQUENCE 324 AA; 34674 MW; 0F1F03927E47A669 CRC64;  
 Query Match 45.1%; Score 723; DB 1; Length 324;  
 Best local similarity 48.7%; Pred. No. 1.2e-50;  
 Matches 149; Conservative 50; Mismatches 93; Indels 14; Gaps 5;

DB 139 VLEGRKDSLTANRSGANDSPFETLAWIPQFNKMGDLIDLVALSGAHTFGARCT 198  
 QY 197 FEGRIY-----DTINASFAALROQTCPSGSDGN-LAPIDVQFVPRDXYFTNLL 248  
 DB 199 FEORLEFNNGSGNPLVTADATFLQTLGICPOGGNNGNTFTNLDISTPNDNDYFTNLQ 258  
 QY 249 SRGFGHSDQELF--NGSGDALVROYASASLFNADVAAMIRGANGVLTAGAGIRR 306  
 DB 259 SNOGLQTDQDELFTSGSNTAIIVRYVGSOTQFFDDVSSMINKGISPLTNGOIRT 318  
 QY 307 NCRVYN 312  
 DB 319 DCKRVN 324  
 RESULT 9  
 PERC\_ARMRU STANDARD; PRT; 332 AA.  
 ID PERC\_ARMRU  
 AC P15233;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Peroxidase C1C precursor (EC 1.11.1.7) (Fragment).  
 DE PRXC1C OR HPRC3.  
 OS Armoracia rusticana (Horseradish) (Armoracia lappatfolia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustosids II; Brassicales; Brassicaceae; Armoracia.  
 NCBI\_TaxID=3704;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88225087; PubMed=3371352;  
 RA Fujiyama K., Takemura H., Shibayama S., Kobayashi K., Choi J.K.,  
 RA Shinyo A., Takano M., Yamada Y., Okada H.;  
 RL "Structure of the horseradish peroxidase Isozyme C genes."  
 CC Eur. J. Biochem. 173:681-687(1988).  
 CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,  
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD  
 CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE  
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.  
 CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
 CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M60729; AAA33379.1; -  
 DR PIR: S00627; S00627.  
 DR HSSP: P00433; 2ATY.  
 DR InterPro: IPR002016; Peroxidase.  
 DR Pfam: PF00141; peroxidase; 1.  
 DR PROSITE: PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 KM Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;  
 KW Multigene family; Signal.  
 FT NON\_TER 1  
 FT SIGNAL <1 9  
 FT CHAIN 10 332  
 FT MOD\_RES 10 10  
 FT ACT\_SITE 47 47 PEROXIDASE C1C.  
 FT ACT\_SITE 51 51 PYRROLIDONE CARBOXYLIC ACID (BY  
 FT METAL 179 179 SIMILARITY).  
 FT DISULFID 20 100 BY SIMILARITY.  
 FT DISULFID 53 58 BY SIMILARITY.



FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .).  
 DR CARBOHYD 298 298 N-LINKED (GLCNAC. . .).  
 SQ SEQUENCE 353 AA; 38825 MW; AC916C03CA24027 CRC64;  
 Query Match 44.1%; Score 707; DB 1; Length 353;  
 Best Local Similarity 46.4%; Pred. No. 2,6e-49;  
 Matches 154; Conservative 40; Mismatches 118; Indels 20; Gaps 5;

QY 2 ASPTLMOCIVAVSLSCVA-----AQLSPTEFVASCSPNIGSIVRAMTQAVASEORMGA 56  
 DB 5 SSSPLFCITILIPVCLLHASLSDAQLTPTFYNSCPNVSNTVRODTIVELMSDPRIIA 64  
 QY 57 SLLELFHDFCVGCGDSILD---AGKRTAGPNLSVSGFEVIDTIRNVEACPGV 112  
 DB 65 SILLELFHDFCVGCGDSILDNTTSFTEKDAFGNANSARGFVIDRMKAIVESACPRF 124  
 QY 113 VSCADILALAAAGDTNLGGPTMSVPLGRDSTTASANSNPPPTASLGLISLFG 172  
 DB 125 VSCADILALAAAGDTNLGGPTMSVPLGRDSTTASANSNPPPTASLGLISLFG 184  
 QY 173 QGIS-PRDMTALSGAHTIGARCTTFRGRIY-----DTDINAFALROOCPRSG 224  
 DB 185 VGLRRSDILVALSGHTFGKNCGRFIMDRLYNFSNTGLPDTLNTTYLQTLRGLCPNGN 244  
 QY 225 DGNLAPIDVQTPVRFPTAFYTNLSRGLFHSDOELF---NGSODALVROYASASL 281  
 DB 245 LSAVLDFDLRTPTFEDKRYVNLKEQGLIOSDOELFSSPNATDTPLVRSFANSTQTEF 304  
 QY 282 ADEVAAMIRMGVGLTGTAGQIRRCRYVNS 313  
 DB 305 NAFVEAMDRGNITPLTGTGQIRLNCRYVNS 336

RESULT 11  
 PERC\_ARMRU STANDARD; PRT: 351 AA.  
 AC P15232;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Peroxidase C1b precursor (EC 1.11.1.7).  
 GN PRXC1B OR HPRC1.  
 OS Amoracia rusticana (Horseadish) (Armoracia laphatifolia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Armoracia.  
 NCBI\_TaxID=3704;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88225087; PubMed=3371352;  
 RA Fujiyama K., Takemura H., Shibayama S., Kobayashi K., Choi J.K.,  
 RA Shimizu A., Takano M., Yamada Y., Okada H.;  
 RT Structure of the horseradish peroxidase isozyme C genes.";  
 RL Eur. J. Biochem. 173:681-687(1988).  
 CC -1- FUNCTION: REMOVAL OF H(2)O(2). OXIDATION OF TOXIC REDUCTANTS.  
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN. DEFENSE RESPONSE TOWARD  
 CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE  
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.  
 CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
 CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M37157; AAA33378.1; .  
 DR PIR: S00626; S00626.

DR HSP: P00433; 2ATJ.  
 DR InterPro: IPR02016; Peroxidase.  
 DR Pfam: PF00141; Peroxidase; 1.  
 DR PRINTS: PR00458; PEROXIDASE. 1.  
 DR PROSITE: PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 DR Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;  
 KW Multigene family; Signal.  
 FT SIGNAL 1 28  
 FT CHAIN 29 351  
 FT MOD\_RES 29 29  
 FT ACT\_SITE 66 66 PEROXIDASE C1B.  
 FT ACT\_SITE 70 70 PYRROLIDONE CARBOXYLIC ACID (BY  
 FT METAL 198 198 SIMILARITY).  
 FT DISULFID 39 119 BY SIMILARITY.  
 FT DISULFID 72 77 IRON (PROTOME IX AXIAL LIGAND).  
 FT DISULFID 125 329 BY SIMILARITY.  
 FT DISULFID 205 237 BY SIMILARITY.  
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 351 AA; 38645 MW; 7A8C606A3928950B CRC64;

Query Match 43.8%; Score 702.5; DB 1; Length 351;  
 Best Local Similarity 46.1%; Pred. No. 5,8e-49;  
 Matches 154; Conservative 38; Mismatches 121; Indels 21; Gaps 5;

QY 1 MASPTLMOCIVAVSLSCVA-----HAQLSPTEFVASCSPNIGSIVRAMTQAVASEORM 54  
 DB 1 MHSPESTFTWTLITLGLAFASLSDAQLTPTFYNSCPNVSNTVRODTIVELMSDPRI 60  
 QY 55 GASLIRLFHDFCVGCGDSILDAG---GEKTAGPNLSVSGFEVIDTIRNVEACPG 110  
 DB 61 TASLIRLFHDFCVGCGDSILDNNTSFLTEKALGNANSARGFVIDRIKAAVERACP 120  
 QY 111 GVSCADILALAAAGDTNLGGPTMSVPLGRDSTTASANSNPPPTASLGLISL 170  
 DB 121 RVSCADILALAAAGDTNLGGPTMSVPLGRDSTTASANSNPPPTASLGLISL 180  
 QY 171 GQGIS-PRDMTALSGAHTIGARCTTFRGRIY-----DTDINAFALROOCPR 222  
 DB 181 AKVGLDRSDILVALSGHTFGKNCGRFIMDRLYNFSNTGLPDTLNTTYLQTLRGLCP 240  
 QY 223 GGDNLAPIDVQTPVRFPTAFYTNLSRGLFHSDOELF---NGSODALVROYASASL 279  
 DB 241 GNQSVLDFDLRTPTFVFNKRYVNLKEQGLIOSDOELFSSPNATDTPLVRSFADG 300  
 QY 280 FNADEVAAMIRMGVGLTGTAGQIRRCRYVNS 313  
 DB 301 FENVEAMDRGNITPLTGTGQIRLNCRYVNS 334

RESULT 12  
 PERC\_ARMRU STANDARD; PRT: 354 AA.  
 AC P24101;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Neuratal peroxidase C precursor (EC 1.11.1.7).  
 GN PRXCA OR AT3G49110 OR T2J13.50.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91200671; PubMed=2016063;



RA Itataprak C., Higashimura N., Yamamoto K., Okada N., Shimmyo A.,  
 RA Takano M.;  
 RT "nucleotide sequences of two genomic DNAs encoding peroxidase of  
 RT Arabidopsis thaliana.";  
 RL Gene 98:237-241(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV, Columbia;  
 RX MEDLINE=21016720; PubMed=11130713;  
 RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unselid M.,  
 RA Partmann B., Valle G., Bloeker H., Perez-Alonso M., Obermayer B.,  
 RA Deleney M., Boutry M., Griwell L.A., Mache R., Pulgomech P.,  
 RA De Simone V., Choise N., Artiguenave F., Robert C., Robert P.,  
 RA Wincker P., Catolico L., Weissenbach J., Saulin W., Queller F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Bernes V.,  
 RA Wurmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,  
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nakamura G.,  
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,  
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,  
 RA Reicheit J., Scharte M., Schoen O., Barques M., Terol J., Clement J.,  
 RA Navarro P., Collado C., Perez-Perez A., Ottenwelder B., Duchemin D.,  
 RA Cooke R., Lande M., Berger-Liauro C., Purnelle B., Masuy D.,  
 RA de Haan M., Maestre A.C., Alcaraz J.-P., Cortet A., Casacuberta E.,  
 RA Monfort A., Argilou A., Flores M., Liguori R., Vitale D.,  
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
 RA Meyer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Creasy T.H., Rizzo M., Walts A., Uteback T., Fujii C.Y., Shea T.P.,  
 RA Creaey T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,  
 RA Pal G., Miltischer J., Sellers P., Gill J.E., Feldblyum T.V.,  
 RA Preuss D., Lin X., Niernan W.C., Salberg S.L., White O., Venter J.C.,  
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asanizu E.,  
 RA Sasamoto S., Kimura T., Idegawa K., Kawashima K., Kishida Y.,  
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shindo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:820-822(2000).  
 CC -1- FUNCTION: REMOVAL OF H(2)O(2). OXIDATION OF TOXIC REDUCTANTS.  
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD  
 CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE  
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.  
 CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
 CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.  
 CC -1- TISSUE SPECIFICITY: ROOTS.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M38380; AAA32849.1;  
 CC EMBL: AL132967; CAB61999.1;  
 CC PIR: J00457; J00457.  
 CC HSP: P00433; 2ATY.  
 CC InterPro: IPR002016; Peroxidase.  
 CC Pfam: PF0014; peroxidase; 1.  
 CC PRINTS: PR00458; PEROXIDASE.  
 CC PROSITE: PS00435; PEROXIDASE\_1; 1.  
 CC PROSITE: PS00436; PEROXIDASE\_2; 1.  
 CC OXidoreductase; Glycoprotein; Peroxidase; Iron; Heme;  
 CC Multigene family; Signal.  
 CC SIGNAL  
 CC FT CHAIN 1 21  
 CC FT ACT\_SITE 22 354 NEUTRAL PEROXIDASE C.  
 CC FT ACT\_SITE 69 69 BY SIMILARITY.  
 CC FT ACT\_SITE 73 73 DISTAL HISTIDINE (BY SIMILARITY).  
 CC FT METAL 201 201 IRON (PROTOME IX AXIAL LIGAND).  
 CC FT DISULFID 42 122 BY SIMILARITY.

FT DISULFID 75 80 BY SIMILARITY.  
 FT DISULFID 128 332 BY SIMILARITY.  
 FT DISULFID 208 240 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 354 AA; 38941 MW; 6D2EED36111724 CRC64;  
 Query Match 43.0%; Score 690; DB 1; Length 354;  
 Best Local Similarity 47.2%; Pred. No. 5; 8e-48;  
 Matches 151; Conservative 37; Mismatches 114; Indels 18; Gaps 5;  
 OY 9 CLVAVSLSCVAAHNOLEPFYFASCCNPLDSIYVRAATQAVASORMGASLRLEPHDFCV 68  
 DB 21 CLM---LCASRSQKQLPTFTYDTSCTPVYINIVDTYVNRSDPNIASILRLPHDFCV 77  
 OY 69 QGCDGSLD---AGGKTAGPNLNSVGFYDITIKRNVACPGVYSCADILATAAR 124  
 DB 78 NGCDASILLDTTSTFTEKDALGNANSARGFPYIDRMKAVERACPRVSCADMILTAQ 137  
 OY 125 DGNILGGPMVPLGRDSTTASGLANSNPPPPASGTLISFGROGIS-PRDMTAL 183  
 DB 138 QSVTLGGSPKVPGLGRDLSLAFLDLANLNPAPFTLPQKANKVNGLDRESDVAL 197  
 OY 184 SGAMTIGQARCTPRGRNYG-----DTDINASFALROOTCPRSGDGNLAPIDVOTP 236  
 DB 198 SGATFGKNGCRFMIDRLYNESNTGILPDLNTTYLQTLRGCCPRNGQSVLVDFDLTP 257  
 OY 237 VRFDTAFNTLSRGLFHSDDOLF---NGSQDALVRYXSASLAFNADFAAIIIRGCGN 293  
 DB 258 LVPDKYVNLKEKGLQSDQELFSSPNATDTIPLVRAVAGCTGTFFNAFEVAMRMGN 317  
 OY 294 VGVLGTAGQIRRCRYVNS 313  
 DB 318 ITPTTGQIRLNCRYVNS 337  
 RESULT 13  
 PERE\_ARATH STANDARD; PRT; 349 AA.  
 AC P24102; O80913;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Basic peroxidase E precursor (Ec 1.11.1.7) (Class III peroxidase  
 DE ATPEA).  
 GN PRXA OR AT2638380 OR T19C21.13.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosids; II: Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA MEDLINE=91200671; PubMed=2016063;  
 RA Itataprak C., Higashimura N., Yamamoto K., Okada N., Shimmyo A.,  
 RA Takano M.;  
 RT "Nucleotide sequences of two genomic DNAs encoding peroxidase of  
 RT Arabidopsis thaliana.";  
 RL Gene 98:237-241(1991).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-CV, Columbia; TISSUE=Leaf;  
 RA Duroux L., Welinder K.G.;  
 RT "Structural diversity and transcription of 40 peroxidases from  
 RT Arabidopsis thaliana.";  
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-CV, Columbia;  
 RX MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Romling C.M., Koo H.L.,  
 RA Moffit K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayen L.,  
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,  
 RA Goodman H.M., Somerville C.R., Copenhagen G.P., Press D.,  
 RA Merman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
 RA Venter J.C.,  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RT Nature 402:761-768(1999).  
 RL (4).  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-CV, Columbia;  
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
 RT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the  
 RT SSP consortium (Salk/Stanford/PGEC).";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,  
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD  
 CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE  
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.  
 CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
 CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.  
 CC -1- TISSUE SPECIFICITY: ROOTS.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M58381; AAA32842.1; -;  
 DR EMBL: AF452388; AAL40852.1; -;  
 DR EMBL: AC004683; AAC28766.1; -;  
 DR EMBL: AY059106; AAL15212.1; -;  
 DR EMBL: AY035033; AAK59538.1; -;  
 DR PIR: J00458; J00458.  
 DR HSSP: P00433; 2ATJ.  
 DR InterPro: IPR002016; Peroxidase.  
 DR Pfam: PF00141; Peroxidase, 1.  
 DR PRINTS: PR00458; PEROXIDASE.  
 DR PROSITE: PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 KW Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;  
 KW Multigene family; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 1 349  
 FT ACT\_SITE 67 349 BASIC PEROXIDASE E.  
 FT ACT\_SITE 71 349 BY SIMILARITY.  
 FT METAL 199 71 DISTAL HISTIDINE (BY SIMILARITY).  
 FT DISULFID 40 120 IRON (PROTOHEME IX AXIAL LIGAND).  
 FT DISULFID 73 78 BY SIMILARITY.  
 FT DISULFID 126 329 BY SIMILARITY.  
 FT DISULFID 206 238 BY SIMILARITY.  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 150 150 L -> Y (IN REF. 1).  
 FT CONFLICT 227 227 L -> Y (IN REF. 1).  
 SQ SEQUENCE 349 AA: 38108 MW: CC9A133616BDAD56 CRC64;

Query Match 42.6%; Score 684; DB 1; Length 349;  
 Best Local Similarity 45.5%; Pred. No. 1,7e-47;  
 Matches 150; Conservative 47; Mismatches 113; Indels 20; Gaps 5;  
 3 SPLMGLVAVSLSCV-----AAHQLSPFYASSCPNLSIYRAMTQAVASQRKMA 56

Db 4 SPSPSCAIALIGLLGLLQASNSNAQLRPDPYFETCTCFVYDIIGNIYDELQDPRIA 63  
 Oy 57 SLRLPFDHDFVQCGGSLLDAG-----GKTAGNLNSVGEFYDTIKRNVACPGV 112  
 Db 64 SLRLPFDHDFVQCGGSLLDAG-----GKTAGNLNSVGEFYDTIKRNVACPGV 123  
 Oy 113 VSCADILALARGTNLGGPTWSVPLGRDSTASASLANSNPPTASITLSIFGR 172  
 Db 124 VSCADILALARGTNLGGPTWSVPLGRDSTASASLANSNPPTASITLSIFGR 183  
 Oy 173 QGLS-PRDNTALSGANTIQACCTFRGRTG-----PRDNTALSGANTIQACCTFRGRTG 224  
 Db 184 VGLNRTSDIALVLSGGHFFERACQFVTPRLVNFNGTNSPDSLPYTLVELRLCPONGN 243  
 Oy 225 DGNLAPIDVQTPPRTFAVFTNLSRGLFHSQDELFN--GGSODALVROYASASILEFA 282  
 Db 244 GTVLNFDVYTPADPFSQYTNLRNGKGLIQSDQLFTPGADITPLVNOYSSDMVSFFR 303  
 Oy 283 DFLVAAIRMGNVGLTGAGQIRNCRVYN 312  
 Db 304 AFLDAMIRMGNLPLGTGTGEIRQNCRVYN 333

RESULT 14  
 PER3\_ARMRU STANDARD; PRT: 349 AA.  
 AC P17180;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Peroxidase C3 precursor (EC 1.11.1.7).  
 GN PRK3.  
 OS Amoracia rusticana (Horse radish) (Amoracia lappatfolia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Equisetids II; Brassicales; Brassicaceae; Amoracia.  
 OC NCBI\_TaxID=3704;  
 RN [1]  
 RP MDLINE=90323613; PubMed-2373366;  
 RA Fujiyama K., Takemura H., Shimizu A., Okada H., Takano M.;  
 RT "Genomic DNA structure of two new horse radish peroxidase-encoding  
 RT genes.";  
 RL Gene 89:163-169(1990).  
 CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,  
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD  
 CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE  
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.  
 CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
 CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D90116; BAA14144.1; -;  
 DR PIR: JH0150; JH0150.  
 DR HSSP: P00433; 2ATJ.  
 DR InterPro: IPR002016; Peroxidase.  
 DR Pfam: PF00141; Peroxidase, 1.  
 DR PRINTS: PR00458; PEROXIDASE.  
 DR PROSITE: PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 KW Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;  
 KW Multigene family; Signal.  
 FT SIGNAL 1 29



```

FT CHAIN 30 349 PEROXIDASE C3.
FT MOD_RES 30 30 PYROLIDONE CARBOXYLIC ACID (BY
FT ACT_SITE 67 67 SIMILARITY).
FT ACT_SITE 71 71 BY SIMILARITY.
FT METAL 199 199 DISTAL HISTIDINE (BY SIMILARITY).
FT METAL 199 199 IRON (PROTOME IX AXIAL LIGAND) (BY
FT DISULFID 40 120 SIMILARITY).
FT DISULFID 73 78 BY SIMILARITY.
FT DISULFID 126 329 BY SIMILARITY.
FT DISULFID 206 238 BY SIMILARITY.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 349 AA; 38179 MW; 6563EB380B75EC3 CRC64;

Query Match 42.3%; Score 679; DB 1; Length 349;
Best Local Similarity 45.0%; Pred. No. 4.3e-47;
Matches 149; Conservative 47; Mismatches 115; Indels 20; Gaps 5;

OY 3 SPTLMQCLVAVSLSCV-----AHAOLSPFTFASSCPNLOSTVRAAMQAVASEQRMCA 56
DB 4 SPLTSCSAMGALLISCCLLQASNSMAOLRPDPFYFRTCPVSFNTIGDIYDELRTDPRIA 63
OY 57 SLRLFFHDFVOCGDCSILLDAG---GEKTAGPNLSVYRGEVIDTIKRYNEACPGV 112
DB 64 SLRLHFDHDFVRCGDSILLDNTSFRTEKDAFNANSARGVDIRKMTSLERACPT 123
OY 113 VSCADITALARDGTNLGFTWVPLGRDSTTASASLANSNPPPTASLGTLISFGR 172
DB 124 VSCADVLTIAISOISVLLSGGFWMPVPLGRDVEAFDLNLTALPSFEFLAQLKRAFAD 183
OY 173 QGLS-PRDMTALSAHTIGARCTTFRRTYG-----DTDNASFAALROQTCPRSG 224
DB 184 VGLNRPDSLVALSOGHFGRAQCFVTPRLYNFNGTRPPTIDPIYVQLRALCQONN 243
OY 225 DGNAPIDVQVPRPDTATYFNLSRGLHSDQELFN--GGSODALVROYASASLENA 282
DB 244 GTVLNEDVVPVTFNFDQRYTNLNGRGLQSDDELSTFGADTIPILVNLSSNTPFAF 303
OY 283 DEVAAMIRMGVAVLTGTAGAIRNCRVNS 313
DB 304 AFVDMATMGRLRLPTGTGEIRNCRVNS 334

RESULT 15
PER2_ARMRU STANDARD; PRT; 347 AA.
AC P17179;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peroxidase C2 precursor (EC 1.11.1.7).
GN PRXC2.
OS Armoracia rusticana (Horseradish) (Armoracia lappatifolia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Armoracia.
OC NGL_TaxID=3704;
KN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90323613; PubMed=2373366;
RA Fujiyama K., Takemura H., Shimizu A., Okada H., Takano M.;
RT "Genomic DNA structure of two new horseradish-peroxidase-encoding
RT genes";
RL gene 89:163-169(1990).
CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,
CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD
CC WOUNDING AND METABOLISM OF AUXIN.. THESE FUNCTIONS MIGHT BE
CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.
CC -1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.
CC -1- COFACTOR: Binds 1 protoheme IX and 1 iron(III) ion.
CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE

```

```

CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: D90115; BAA14143.1; -
DR PIR: JH0149; JH0149.
DR HSSP: P00433; 2ATJ.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00141; peroxidase; 1.
DR PRINTS: PR00458; PEROXIDASE.
DR PROSITE: PS00435; PEROXIDASE_1; 1.
DR PROSITE: PS00436; PEROXIDASE_2; 1.
DR Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;
KW Multigene family; Signal.
KM SIGNAL 1 24
FT CHAIN 25 347 PEROXIDASE C2.
FT MOD_RES 25 25 PYROLIDONE CARBOXYLIC ACID (BY
FT ACT_SITE 62 62 SIMILARITY).
FT ACT_SITE 66 66 BY SIMILARITY.
FT METAL 194 194 DISTAL HISTIDINE (BY SIMILARITY).
FT METAL 194 194 IRON (PROTOME IX AXIAL LIGAND) (BY
FT DISULFID 35 115 SIMILARITY).
FT DISULFID 68 73 BY SIMILARITY.
FT DISULFID 121 325 BY SIMILARITY.
FT DISULFID 201 233 BY SIMILARITY.
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 347 AA; 38035 MW; 3EE9A2CEDEBBA9A CRC64;

Query Match 42.2%; Score 676.5; DB 1; Length 347;
Best Local Similarity 47.0%; Pred. No. 6.8e-47;
Matches 151; Conservative 38; Mismatches 109; Indels 23; Gaps 6;

OY 10 LVAVSLSCVAHAOLSPFTFASSCPNLOSTVRAAMQAVASEQRMCA5LRLFFHDFVOC 69
DB 16 LINVSL-----SHAOLSPFTFDTCTPOVFDATVNTITKALSDPRIASLRLHFDHCFV 71
OY 70 GDCGSILLD---AGGKTAGPNLSVYRGEVIDTIKRYNEACPGVSCADITALARD 125
DB 72 GCDASILLDNTSFRTEKDAFNANSARGVDIRKMTSLERACPTVSCADLITALAOK 131
OY 126 GTNLGFTWVPLGRDSTTASASLANSNPPPTASLGTLISFGRGLS-PRDMTALS 184
DB 132 SVVLAGGSMVPSGRDRSLNGFMDLANDNLPGSSSTLYQLKFRVNGIDRSDLYALS 191
OY 185 GAHTIGQARCTTFRRTYG-----DTDNASFAALROQTCPRSGDGTIAPIDVQTV 237
DB 192 GGHFTGKNGQCFIMDRILNFNSGKRPPTIDKSLYSLRKGCRNGNLVLYVDFLRTPT 251
OY 238 RPDATYFTNLNRGLHSDQELFN-----LVROYASASLEFNADFPAAITRMG 292
DB 252 IFDNKYVYVNLKNGGLQSDDELFS--SPDASDTPILVRAVADGGKFEFADFAEAMIRMG 309
OY 293 NVGVLGTAGQAIRNCRVNS 313
DB 310 NLSPTGKQGEIRNCRVNS 330

```

Search completed: April. 26, 2003, 12:34:05  
 Job time : 27 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:29:18 ; Search time 34 seconds

(without alignments)  
1896.845 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 1604

Sequence: 1 MASPTMOCIVAVSLSCVA.....VGVLTGTAQIRNCRVNS 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP archaea:  
2: SP bacteria:  
3: SP fungi:  
4: SP human:  
5: SP invertebrate:  
6: SP mammal:  
7: SP mhc:  
8: SP organelle:  
9: SP phage:  
10: SP plant:  
11: SP rodent:  
12: SP virus:  
13: SP vertebrate:  
14: SP unclassified:  
15: SP virus:  
16: SP bacteriophage:  
17: SP archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1292	80.5	307	10	Q43416
2	1111	69.3	316	10	O9XGV6
3	1004.5	62.6	309	10	P93550
4	997.5	62.2	305	10	O65029
5	982.5	61.1	319	10	O91LM9
6	980	61.1	324	10	O9FLC0
7	972.5	60.6	333	10	P93548
8	958.5	59.8	321	10	O9XIV9
9	950	59.2	318	10	O9SSZ7
10	935	58.3	311	10	O9XKY9
11	935	58.3	315	10	O24336
12	928	57.9	317	10	O43006
13	924	57.6	320	10	O41324
14	922	57.5	314	10	O22441
15	919	57.3	317	10	O22438
16	918	57.2	320	10	O8RVP7

17	915.5	57.1	316	10	Q41577	triticum ae
18	912.5	56.9	325	10	O91VL1	arabidopsis
19	895.5	55.8	319	10	Q43220	triticum ae
20	887.5	55.3	314	10	O22440	oryza sativ
21	877.5	54.7	313	10	Q43417	arabidopsis
22	872	54.4	324	10	O9XFL4	phaseolus v
23	870.5	54.3	316	10	O91VL2	arabidopsis
24	867.5	54.1	315	10	O22439	oryza sativ
25	867	54.1	314	10	Q43218	triticum ae
26	849	52.9	325	10	O9SSZ8	scutellaria
27	848.5	52.9	323	10	Q42904	linum usita
28	846.5	52.6	315	10	P93549	spinaclia ol
29	844.5	52.6	321	10	O49193	striga asia
30	839.5	52.3	313	10	O81524	avena sativ
31	838.5	52.3	312	10	Q43212	triticum ae
32	834.5	52.0	322	10	O49192	striga asia
33	828	51.6	335	10	O91L45	oryza sativ
34	822	51.2	303	10	O91EH6	hordeum vul
35	813.5	50.7	310	10	O9M909	hordeum vul
36	812	50.6	315	10	O40068	arabidopsis
37	811	50.6	314	10	O81525	hordeum vul
38	808	50.4	331	10	O23237	avena sativ
39	801	49.9	315	10	O91E15	arabidopsis
40	801	49.9	338	10	O9S116	arabidopsis
41	799.5	49.8	325	10	O40372	medicago tr
42	798.5	49.8	323	10	O92T57	zea mays (m
43	794.5	49.5	329	10	O94100	zea mays (m
44	790.5	49.3	334	10	O9M422	spinaclia ol
45	787	49.1	345	10	Q43102	populus trl

## ALIGNMENTS

RESULT 1  
Q43416 PRELIMINARY; PRT; 307 AA.

AC 043416; 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)  
DE Peroxidase precursor (EC 1.11.1.7).  
OS Cenchrus ciliaris (Bufolegrass) (Pennisetum ciliare).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Paniceae; Cenchrus.  
OX NCBI\_TaxID=35872;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. BILIOELA; TISSUE=WOONDED AND UNWOONDED LEAVES, AND STEMS;  
RA Ross A.H.;  
RT \*Investigation of peroxidase genes and genetic transformation in  
RT buffel grass.\*  
RL Thesis (1994). Botany, University of Queensland.  
DR EMBL: U12314; AAA20472.1; -.  
DR HSSP: P22195; ISCH.  
DR InterPro: IPR002016; Peroxidase.  
DR Pfam: PF00141; Peroxidase; 1.  
DR PRINTS: PR00458; PEROXIDASE.  
DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
KW Oxidoreductase; Peroxidase; Signal.  
FT SIGNAL 1 23  
FT CHAIN 24 307 POTENTIAL.  
SQ SEQUENCE 307 AA; 32507 MW; 526B44B3A7C75FA7 CRC64;

Query Match 80.5%; Score 1292; DB 10; Length 307;  
Best Local Similarity 81.5%; Pred. No. 2e-96;  
Matches 255; Conservative 24; Mismatches 28; Indels 6; Gaps 4;

QY 1 MASPTMOCIVAVSLSCVAHQSPTFYASSCPNLSIVRAAMQAVASQRMGASLLR 60  
DB 1 MAAPTLMOCIVAVSLSCVAHQSPTFYASSCPNLSIVRAAMQAVASQRMGASLLR 60

not available in  
STC-man

OY	61	LFFHDCFYOGDDGSIILDAGCEKTAQPNLNSVSGFEVIIDITIKRNVEACGVVSCADIIA	120
Db	61	LFFHDCFYOGDDGSIILSRSGEKSAGSNANSVSGIYEVIDTIKKNVEACGVVSCAITIVF	120
OY	121	LAARDGTNLGGPTWASVPLGRDSTTASASLANSPPPPAISGLTSLIFRGÖLSBDM	180
Db	121	LAARPGPMLLGCPWWNPVLGRDSTTMLSLANONLEPPF-SLGTLSLISGGR-LSARM	178
OY	181	TALSGAHFIGOARTTFRGRITYGPTDINASFALROOTCRSGGDGLAPIDVOTPARFD	240
Db	179	IATSGAAHH-AQARKTTFRGRITYGDTINDASFALQQOOTCRSSGSDGLAIIDMOTPARFD	237
OY	241	TAYFTNNLSRRGLESHDOELFNNGSQDALVRQYSASASLENADFVAAMIRMGNVGLTGT	300
Db	238	NAYTNINLVSRRGLEPHSDQELFNNGSQDALVROYSSPSQENSDFAAMIMKGNTIG--AN	294
OY	301	AGOIRBNCRVVNS	313
Db	295	AGOVBRNCRVVNS	307

RESULT 2				
09XGV6				
ID	09XGV6	PRELIMINARY;	PRT;	316 AA.
AC	09XGV6;			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, last annotation update)			
DE	Bacterial-induced peroxidase precursor (EC 1.11.1.7).			
OS	Gossypium hirsutum (Upland cotton).			
OC	Eumariota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;			
CC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;			
CC	eucosids. II: Malvales; Malvaceae: Gossypium.			
OX	NCBI_TaxID:3635;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV. RBDA B50.			
RA	Assiglaette K., Cuny G., Valette C., Delannoy E., Bresson E.,			
RA	Jallouli A., Daniel J.-F., Geiger J.-P., Nicole M.;			
RT	*Cloning and Characterization of a Bacterial-Induced Peroxidase-			
RL	Encoding cDNM from Cotton (Accession No. AF155124). (PGR99-129).";			
DR	EMBL; AF155124; AAd03561.1; -.			
DR	HSSP; P22195; 1SCH.			
DR	InterPro; IPR002016; Peroxidase.			
DR	Pfam; PF00141; peroxidase; 1.			
DR	PRINTS; PS00458; PEROXIDASE.			
DR	PROSITE; PS00435; PEROXIDASE.1; UNKNOWN_1.			
DR	PROSITE; PS00435; PEROXIDASE.2; 1.			
ET	Oxidoreductase; Peroxidase; Signal.			
FT	SIGNAL			
FT	CHAIN			
FT	SEQUENCE			
30	316 AA; 33696 MW; 8EBIC05BDEC47P96 CRC64;			
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				

Query Match	69.3%;	Score 1111;	DB 10;	Length 316;
Best Local Similarity	69.7%;	Pred. No. 8.5e-82;		
Matches 221; Conservative	30;	Mismatches 60;	Indels 6;	Gaps 3

[illegible]

Qy 236 PVDEPTAETMLNLRGGLFHSDDOLENGSODALVROYSA\$ASL\$F\$N\$D\$E\$F\$V\$A\$M\$M\$N\$V\$G 295

Db 240 PTFRENDNYFRNLDVARRKGLHSDOLEFNGSODALVRYT\$NN\$P\$A\$F\$S\$A\$D\$F\$A\$A\$A\$W\$K\$M\$N\$IS 299

Qy 296 VLTGAGQIRNRCRYVN 312

Db 300 PLTGTGGEIRNRCRYVN 316

### RESULT 3

ID	P93550	PRELIMINARY;	PRT;	309	AA
----	--------	--------------	------	-----	----

DT 01-MAY-1997 (TREMBLrel. 03, Created,

DT 01-MAR-2002 (Tremblay, 20, Last annotation update)

PRXR6. GN

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta,

Caryophyllales; Chenopodiaceae; Splanclia.

RN [1]

RC TISSUE=LEAF;

"Diversity and conservation of plant peroxidases." *Plant Peroxidase Newsletter* 1:4-7(1993)

RN [2] EREFERENCE FROM N A

RC TISSUE=LEAF;  
PA cimon D ;

RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases

DR HSSP; PZL195; LSCH.  
DB InterPro: TPR002016: Peroxidase

DR Pfam: PF00141: peroxidase\_1

PROSITE: PS00435: PEROXIDASE

DR PROSITE: PS01331: THYMIDYLATE KINASE: 1.  
DR PROSITE: PS00430, ENOYLAIDASE\_2, 1.

FT	NON TER	1	1
AM	OXIDOLCUCURCASE, FEROXALCUCUR, DIGNAT:		

FT	CHAIN	13	309	POTENTIAL.
F1	STONAL	21	12	POTENTIAL.

SEQUENCE	NO. FM,	NO. OF FM,	CODED/DELETED/NO
30			

Query Match	62.6%	Score 1004.5	DB 10	Length 309
Best Local Similarity	63.1%	Pred No. 3.3e-73		
Matches 195	Conservative 46	Mismatches 61	Indels 7	Gaps 3

Qy 11 VAVSLISCVAHAQISPTFEAASCCNLOSIVRAAMTQAVASEORNGASILRLFFEDHCFOG 70  
::  
Db 1 ITIATLACLSNAOISSKHYAASCNCNLKITYRKTKMQAYQEKQRNGASILRLFFEDHCFVNG 60

QY	71	CDGSLLD---	AGGEKTGPNL	NSVRFEEVDIT	KRVNEAC	PGVSCADIL	LAARD	125
Db	61	CDASLLD	STFGKTAIS	NNRNSV	RGEVDIS	IKTVNEAS	KATVSCADIL	LAARD 120

Qy 126 GTNLLGGPWPVPLGRDSTTASASLANSPNPPTASTLGLSLFGRQLSPKMTALSG 185  
+ ||||| : : : : :  
Db 121 GVFLGGSWVKPLGRDRATASTAATNNLPASSSSINTLTFFNNKKISPKMTALSG 180

[illegible]

244 FTNLSRGLFHSDOELFNGSGODALVROYASASLFNADFEVAMIRMGVGLGTACQ 303

OY 304 IRRNCRVYN 312  
 11:111 :1  
 Db 301 IRRNCRFIN 309

## RESULT 4

065029 ID 065029 PRELIMINARY; PRT; 305 AA.  
 AC 065029;  
 DT 01-AUG-1998 (TEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
 DE Peroxisome FLXPER4 (EC 1.11.1.7) (fragment).  
 GN PER4.  
 OS Linum usitatissimum (Flax) (linseed).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Linaceae; Linum.  
 NCBI\_TaxID=4006;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. STORMONT CIRROS; TISSUE=LEAF;  
 RA Oman F., Tyson H.;  
 RT "CDNA sequence of a basic peroxidase (FLXPER4) in flax.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF049881; AAC05277.1; -.  
 DR HSSP; P22195; ISCH.  
 DR InterPro; IPR002016; Peroxidase.  
 DR Pfam; PF00141; peroxidase; 1.  
 DR PRINTS; PR00458; PEROXIDASE.  
 DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
 DR PROSITE; PS00436; PEROXIDASE\_2; 1.  
 DR Oxidoreductase; Peroxidase.  
 KW NON\_TER  
 FT SEQUENCE 305 AA; 32431 MW; F79BBIAD697253 CRC64;  
 SQ

Query Match 62.2%; Score 997.5; DB 10; Length 305;  
 Best local similarity 64.6%; Pred. No. 1.2e-72;  
 Matches 197; Conservative 37; Mismatches 66; Indels 5; Gaps 2;

OY 12 AVSLSCVAHAQISPTFYASCPNLIQSTIVRAAMQAVASEQRMGASLIRLFHDCFYOGC 71  
 11:111 :1  
 Db 2 AMSLSSGSAQLANFATSCPTLLITVRNMAQVANSERMAASIRLHFHCFVNGC 61  
 OY 72 DGSILLD----AGGKTKAPNINSVRGFEVIDTIRNVEACPGVSCADILALAAARGT 127  
 11:111 :1  
 Db 62 DGSILLDATATFTGERKMGPNQNSVGFDDIDITIRVEACMATVSCADILALAAARGV 121  
 OY 128 NLGGPTVMSVPLGRDSTTASASLANSNPPPTASLGLTSLFGRQGLSPRDMTALSGA 187  
 11:111 :1  
 Db 122 VLVGPTVMSVPLGRDARTASASLANSNPPPTASLGLTSLFGRQGLSPRDMTALSGA 181  
 OY 188 TTGOARCTTFRGRITVDIDINAFALRQOTCPRGSGGONLAPIDVOTPVREPTATYFTNL 247  
 11:111 :1  
 Db 182 TTGOARCTTFRGRITVDIDINAFALRQOTCPRGSGGONLAPIDVOTPVREPTATYFTNL 240  
 OY 248 LSRGGLFHSDDQLFNGSGSDALVROYASASLNFADVAAMIRMGNGVLTGTACQIRRN 307  
 11:111 :1  
 Db 241 VARRGLHSDQLFNGSGSDALVROYASASLNFADVAAMIRMGNGVLTGTACQIRRN 300  
 OY 308 CRVYN 312  
 11:111 :1  
 Db 301 CRPN 305

RESULT 5  
 09LMM9 ID 09LMM9 PRELIMINARY; PRT; 319 AA.  
 AC 09LMM9;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)

DE ESTS D41670(54325).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euphorbiaceae; Oryzae; Oryza.  
 NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsunoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC  
 clone: P0451C06.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP001551; BAA92967.1; -.  
 DR HSSP; P22195; ISCH.  
 DR InterPro; IPR002016; Peroxidase.  
 DR Pfam; PF00141; peroxidase; 1.  
 DR PRINTS; PR00458; PEROXIDASE.  
 DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
 DR PROSITE; PS00436; PEROXIDASE\_2; 1.  
 DR SEQUENCE 319 AA; 33536 MW; 395ELBE20BDEB119 CRC64;  
 SQ

Query Match 61.3%; Score 982.5; DB 10; Length 319;  
 Best local similarity 61.5%; Pred. No. 2e-71;  
 Matches 192; Conservative 42; Mismatches 69; Indels 9; Gaps 3;

OY 11 VAVSLSCVAHAQISPTFYASCPNLIQSTIVRAAMQAVASEQRMGASLIRLFHDCFYOGC 70  
 11:111 :1  
 Db 7 VALAVFLAEAOQLSPGYVATCPGVASIVRRMAQAVOKESMGASLIRLFHCFVNGC 66  
 OY 71 CDGSILLD----AGGKTKAPNINSVRGFEVIDTIRNVEACPGVSCADILALAAARGT 126  
 11:111 :1  
 Db 67 CDASILLDPTANFTGERKMGPNQNSVGFDDIDITIRVEACMATVSCADILALAAARGV 126  
 OY 127 TNLGGPTVMSVPLGRDSTTASASLANSNPPPTASLGLTSLFGRQGLSPRDMTALSGA 186  
 11:111 :1  
 Db 127 VNLGGPTVMSVPLGRDARTASASLANSNPPPTASLGLTSLFGRQGLSPRDMTALSGA 186  
 OY 187 HTGOARCTTFRGRITVDIDINAFALRQOTCPRGSGGONLAPIDVOTPVREPTATYFTNL 245  
 11:111 :1  
 Db 187 HTVGMAWCTTFRGRITVDIDINAFALRQOTCPRGSGGONLAPIDVOTPVREPTATYFTNL 246  
 OY 246 NLSRGLFHSDDQLFNGSGSDALVROYASASLNFADVAAMIRMGNGVLTGTACQIRRN 301  
 11:111 :1  
 Db 247 DLSRGLFHSDDQLFNGSGSDALVROYASASLNFADVAAMIRMGNGVLTGTACQIRRN 306  
 OY 302 GOIRNCRVYN 313  
 11:111 :1  
 Db 307 GEVAINCRVYN 318

RESULT 6  
 09FLCO ID 09FLCO PRELIMINARY; PRT; 324 AA.  
 AC 09FLCO;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)  
 DE Peroxidase (putative peroxidase).  
 GN AT5G05340.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=98344145; Pubmed=9679202;  
 RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,  
 Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence  
 features of the regions of 1,381,565 bp covered by twenty one

RT physically assigned pl and TAC clones.;"

RL DNA Res. 5:131-145(1998).

RN [12]

RP SEQUENCE FROM N.A.

RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Meyers M., Miranda M., Seki M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Shim P., Southwick A., Shinzaki K., Davis R.W., Eckert J.R., Theologis A.

RA "Full Length cDNA of gene At5g05340 (GI:15239075).";

RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB010692; BAB09977.1; -

DR EMBL; AY065270; AAL38746.1; -

DR HSSP; P22195; ISCH.

DR InterPro: IPR002016; Peroxidase.

DR Pfam: PF00141; peroxidase; 1.

DR PRINTS; PR00458; PEROXIDASE.

DR PROSITE; PS00435; PEROXIDASE\_1; 1.

DR PROSITE; PS00436; PEROXIDASE\_2; 1.

DR Peroxidase.

SW SEQUENCE 324 AA; 34215 MW; F891CD4EEBEC5AD4 CRC64;

Query Match 61.1%; Score 980; DB 10; Length 324;

Best Local Similarity 61.4%; Pred. No. 3.3e-71;

Matches 199; Conservative 36; Mismatches 77; Indels 12; Gaps 3;

QY 1 MASPTLMQCLVAVSL-----SCVAHQLSPTFYAASCPNLSIVRAAMTQAVASEQRM 54

DB 1 MASNKLTSLVLYVLLILGDNVNYEVLQITNFYSTSCPNLSTVQTVASVASEARM 60

QY 55 GASLRFHDFCVOCDSIILD---AGGKTAGPNLNSVGRFEVDTIKRNEACP 110

DB 61 GASILRFHDFCVOCDSIILD---AGGKTAGPNLNSVGRFEVDTIKRNEACP 120

QY 111 GVSCADITALAARDGTMLGPTWSVPLGRDSTTASASIANSPPTASLGLTSLF 170

DB 121 GVSCADITALAARDGTMLGPTWSVPLGRDSTTASASIANSPPTASLGLTSLF 180

QY 171 GGGISPRDMALSGAHTTGAARCTTFRGRTYGTDTINASFALRQOTCPSPS--GGDGNL 228

DB 181 SAVGLSTRDMVALSGAHTTGAARCTTFRGRTYGTDTINASFALRQOTCPSPS--GGDGNL 240

QY 229 APIDVQTVPRFDYAFYTNLSRGLFHSDELFGSGSDALVROYASASLFNADFEVAA 288

DB 241 APIDVQTVPRFDYAFYTNLSRGLFHSDELFGSGSDALVROYASASLFNADFEVAA 300

QY 289 IRMGNGVLTGTAGQIRNRCRVN 312

DB 301 IRMGNGVLTGTAGQIRNRCRVN 324

RESULT 7

P33548 PRELIMINARY; PRT; 323 AA.

AC P33548;

DT 01-MAY-1997 (Tremblrel. 03, Created)

DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Peroxidase precursor (EC 1.11.1.7).

GN PRXR4.

OS Spinacia oleracea (Spinach).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.

OC NCBI\_TaxID=3362;

OX [1]

RN SEQUENCE FROM N.A.

RP TISSUE-LEAF;

RA Simon P.;

DR Pfam: PF00141; peroxidase; 1.

DR PRINTS; PR00458; PEROXIDASE.

RL Plant Peroxidase Newsletter 1:4-7(1993).

RN [12]

RP SEQUENCE FROM N.A.

RA Simon P.;

RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; Y10465; CAA71491.1; -

DR HSSP; P22195; ISCH.

DR InterPro: IPR002016; Peroxidase.

DR Pfam: PF00141; peroxidase; 1.

DR PRINTS; PR00458; PEROXIDASE.

DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.

DR PROSITE; PS00436; PEROXIDASE\_2; 1.

DR Oxidoreductase; Peroxidase; Signal.

FT SIGNAL 27

FT CHAIN 28 323

FT POTENTIAL.

SW SEQUENCE 323 AA; 35283 MW; 8547E125CDS6E75E CRC64;

Query Match 60.6%; Score 972.5; DB 10; Length 323;

Best Local Similarity 61.0%; Pred. No. 1.3e-70;

Matches 197; Conservative 39; Mismatches 76; Indels 11; Gaps 4;

QY 1 MASPTLMQCLVAVSLSCVAHQ-----LSPTFYAASCPNLSIVRAAMTQAVASEQRMGA 56

DB 1 MASMLVVGFLAIFSLILLAGTSDAMLKRPFYAASCPNLSIVRAAMTQAVASEQRMGA 60

QY 57 SLILFHFDFCVOCDSIILD---AGGKTAGPNL-NSVGRFEVDTIKRNEACPG 111

DB 61 SLILFHFDFCVOCDSIILD---AGGKTAGPNL-NSVGRFEVDTIKRNEACPG 120

QY 112 VSCADITALAARDGTMLGPTWSVPLGRDSTTASASIANSPPTASLGLTSLF 171

DB 121 TVSCADITALAARDGTMLGPTWSVPLGRDSTTASASIANSPPTASLGLTSLF 180

QY 172 RQGLSPDMALSGAHTTGAARCTTFRGRTYGTDTINASFALRQOTCPSPS--GGA 229

DB 181 RQGLSPDMALSGAHTTGAARCTTFRGRTYGTDTINASFALRQOTCPSPS--GGA 240

QY 230 PIDVQTVPRFDYAFYTNLSRGLFHSDELFGSGSDALVROYASASLFNADFEVAA 289

DB 241 PIDVQTVPRFDYAFYTNLSRGLFHSDELFGSGSDALVROYASASLFNADFEVAA 300

QY 290 IRMGNGVLTGTAGQIRNRCRVN 312

DB 301 IRMGNGVLTGTAGQIRNRCRVN 323

RESULT 8

O9XIV9 PRELIMINARY; PRT; 321 AA.

AC O9XIV9;

DT 01-NOV-1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE Peroxidase (EC 1.11.1.7).

OS Nicotiana tabacum (Common tobacco).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanales; Solanales; Nicotiana.

OC NCBI\_TaxID=4097;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=CV. SAMSON NN; TISSUE=TWY-INFECTED LEAF;

RA MEDLINE=99418329; PubMed=10490396;

RA Hiraoka S., Ito H., Matsui H., Honma M., Ohashi Y.;

RT "cDNA sequences for two novel tobacco peroxidase isoenzymes (Accession Nos. AB027752 and AB027753) (PGR99-109).";

RT Plant Physiol. 120:1205-1205(1999).

DR EMBL; AB027752; BAA82306.1; -

DR HSSP; P22195; ISCH.

DR InterPro: IPR002016; Peroxidase.

DR Pfam: PF00141; peroxidase; 1.

DR PRINTS; PR00458; PEROXIDASE.

Db 239 NAYYTNLSNKGLLHSDYLFENNGSGIDNTYVNFASNAAFSSAFATAMVMNMGINAPKGT 298

QY 301 AGOIRNCRVNS 313  
 DB 299 NGOIRLSCSVNS 311

## RESULT 11

024336 PRELIMINARY; PRT: 315 AA.

AC 024336: 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Korean-radiash isoperoxidase (EC 1.11.1.7).  
 GN PRXK1.  
 OS Raphanus sativus (Radish).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosids II; Brassicales; Brassicaceae; Raphanus.  
 OX NCBI\_TaxID=3726;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. HANDSOME HALL;  
 RA Jong Hoon P.;  
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.  
 RN (12)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. HANDSOME HALL;  
 RA Park J.;  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: X51172; CAA62597.1; -.  
 DR HSSP: P22195; ISCH.  
 DR InterPro: IPR002016; Peroxidase.  
 DR Pfam: PF00141; peroxidase; 1.  
 DR PRINTS: PS00435; PEROXIDASE.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 DR OXidoreductase; Peroxidase.  
 KW SEQUENCE 315 AA; 33558 MW; B6F45C53EB845523A CRC64;

Query Match 58.3%; Score 935; DB 10; Length 315;  
 Best Local Similarity 59.9%; Pred. No. 1,4e-67;  
 Matches 184; Conservative 44; Mismatches 75; Indels 4; Gaps 2;

QY 10 LVAVSLSCVNAHQSPFTYASCPNLOSIVRAAMTQAVASEQMGSLRFEFHDFVQ 69  
 DB 9 VIVITLLGGEOITFTYSTSCPNLLSTYSGVKSAYSSQPMGASIRLFFHDFV 68  
 QY 70 GCDGSITLD--AGEKTAGPMLNSVGRFEVIDTKRNVFAACGVVSCADIIAARDG 126  
 DB 69 GCGSITLDDSTFTGQNGAPNRNSARGFVINDIKSAVEACGVVSCADIIAARDG 128  
 QY 127 TNLGGPTWSVPLGRDSTTASASLANSNPPPTASLGTLSLFGROGLSPDMTALSGA 186  
 DB 129 VVVLGGPNNVVKVGRDASTASQAANSNIPASWSLSQLISPRVGLSTRDMVALSGA 188  
 QY 187 HTTGOARCTFGRGRTIGDIDINAFALROOCPRGSG--GDGNLAPIVOTVPRFDATYFF 245  
 DB 189 HTTGSRCTNFRRTIYNENINAFATLRKSCPPAARRKKPPDLINSPSTFSDNSYK 248  
 QY 246 NLSRGRGLFHSDOELFNGSGODALVROYASASLFNADFAAMTGMNGVGLTGTACQIR 305  
 DB 249 NLAAQGLLHSDQVLEFNGSGSTSIYRGYSNPSFNSDFAAAMTKMDISPLTSGSGEIR 308  
 QY 306 RNCRVVN 312  
 DB 309 KVCGRTN 315

## RESULT 12

043006 PRELIMINARY; PRT: 317 AA.

AC 043006: 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Peroxidase precursor (EC 1.11.1.7).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. NOHRIN;  
 RA Reihmann C., Ringli C., Duder R.;  
 RT Complementary DNA cloning and sequencing of a pathogen-induced  
 RT putative peroxidase from rice."  
 RL Plant Physiol. 100:1611-1612(1992).  
 DR EMBL: X66125; CAA46916.1; -.  
 DR HSSP: P22195; ISCH.  
 DR InterPro: IPR002016; Peroxidase.  
 DR Pfam: PF00141; peroxidase; 1.  
 DR PRINTS: PS00435; PEROXIDASE.  
 DR PROSITE: PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE: PS00436; PEROXIDASE\_2; 1.  
 KW OXidoreductase; Peroxidase; Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 26 317  
 FT POTENTIAL. PEROXIDASE.  
 SQ SEQUENCE 317 AA; 32875 MW; CEC727D0DA30E311 CRC64;

Query Match 57.9%; Score 928; DB 10; Length 317;  
 Best Local Similarity 58.5%; Pred. No. 5e-67;  
 Matches 186; Conservative 44; Mismatches 82; Indels 6; Gaps 3;

QY 1 MASPT---LMOCLVAVSLSCVNAHQSPFTYASCPNLOSIVRAAMTQAVASEQMGAS 57  
 DB 1 MASATNSLSIMLVAAAMSVASAOISATFYDTSCPNALSTIKSVITAAVNSERRGAS 60  
 QY 58 LRLRFHDFCFVCGCDASVLL--SGQEQNAGPNVSGFSEVIDTKRNVFAACPGVSCAD 117  
 DB 61 LRLRFHDFCFVCGCDASVLL--SGQEQNAGPNVSGFSEVIDTKRNVFAACPGVSCAD 119  
 QY 118 ILAALARDGTLNLGCEPTWSVPLGRDSTTASASLANSNPPPTASLGTLSLFGROGLSP 177  
 DB 120 ILAVALARDSVVALGSPSWVTLGRDSTTASBALANTDLPAPSSSLMELIGNFSRKLDA 179  
 QY 178 RDMALSGATTGACCTFGRGRTIGDIDINAFALROOCPRG--SGGNSNLAPIVOT 235  
 DB 180 TDVVALSGATTGACCTFGRGRTIGDIDINAFALROOCPRG--SGGNSNLAPIVOT 239  
 QY 236 PVRFDTAYFTNLSRGLFHSDOELFNGSGODALVROYASASLFNADFAAMTGMNGV 295  
 DB 240 PNAFDNAYVNSNLSNGLHSDQVLEFNGSGADNVTRNFASNAARSAFTTAAYKMGNIS 299  
 QY 296 VLTGTACQIRNCRVNS 313  
 DB 300 PLTGTQOIRLSCSVNS 317

## RESULT 13

041324 PRELIMINARY; PRT: 320 AA.

AC 041324: 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Cationic peroxidase.  
 OS Stylisanthes humilis (Tomswayille stylio).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Stylisanthes.  
 OX NCBI\_TaxID=35628;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PATERSON; TISSUE-LEAF;





Wed May 7 14:10:03 2003

us-10-047-825-4.rspt

Page 8

```

QY 118 ILLAROSTNLLGGPMSVPLGRRROSTTASLASNPNPPATSLGILLSPRQGLSP 177
Db 120 ILLAARSDVAALGSPSPMTVLLGRROSTTASLAAITDILPAVSSSLAELINFERRKGLDA 179

QY 178 RDMTALSGAHTIGQARCTTFGRGYGDTDNASFAALROOTCPR--SGGDDNLAIPDIVOT 235
Db 180 TDMVALTSGAHTIGQAQCNCFRDRYNTNNTIDSAFQIRORANPRRTSGSDSLAAADPTT 239

QY 236 PVRFTAYATNTLISRRGLFHSQDELFNNGSQDALVROYSASASLFNADFFAAMTRKGNVG 295
Db 240 PNAFDNAYYSNLLSKNGLLHSDOVLFNNGSGADNTVNFASNAAPSAFTTAAVKKMNIS 299

QY 296 VLTGTAGAIRRCRRVCN 313
Db 300 PLTGIGQIRLSCSRVNS 317

```

Search completed: April 26, 2003, 12:34:46  
Job time : 37 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 3, 2003, 11:28:57 ; Search time 2153 seconds  
(without alignments)  
4230.923 Million cell updates/sec

Title: US-10-047-825-4  
Perfect score: 1604  
Sequence: 1 MASPTLMOCIVAVSLSCVA.....VGVLTGTAGQIRNRCRVNS 313

Scoring table: BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.US10047825/runat.26042003.111958.7132/app\_query.fasta.1.455  
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US10047825.@cgn2.1.1.1687@runat.26042003.111958.7132 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMECUT=120  
-WARN\_TIMECUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXt=7

Database: GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_com:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_jmu:\*  
20: em\_com:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_nam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1292	80.5	1335	8	CCU12314	U12314 Genchirus ci
2	1111	69.3	1369	8	AF155124	AF155124 Gossypium
3	1040.5	64.9	160562	2	AP005613	AP005613 Oryza sat
4	1036.5	64.6	135792	2	AP003544	AP003544 Oryza sat
5	1036.5	64.6	158456	2	AP004731	AP004731 Oryza sat
6	1004.5	62.6	1091	8	SOPRKR6	Y10467 S.oleracea
7	997.5	62.2	1176	8	AF049881	AF049881 Linum usi
8	980	61.1	1238	8	AY065270	AY065270 Arabidops
9	972.5	60.6	1176	8	SOPRKR4	Y10465 S.oleracea
10	958.5	59.8	1356	8	AB027752	AB027752 Nicotiana
11	950	59.2	957	8	AB024439	AB024439 Scutellar
12	936.5	58.4	1144	8	SSNPEROXIA	L36110 Stylosanthe
13	935	58.3	1330	8	AF247700	AF247700 Oryza sat
14	928	57.9	1287	8	OSPER	X66125 O.sativa mr
15	927	57.8	1116	8	ARCPNCL	M7636 Arabidops hyp
16	926	57.7	1310	8	RICPERX	D16447 Oryza sativ
17	919.5	57.3	151773	8	AF001551	AF001551 Oryza sat
18	919	57.3	1306	8	AF014467	AF014467 Oryza sat
19	918	57.2	1336	8	AF485265	AF485265 Gossypium
20	896.5	55.9	1360	8	AF037034	AF037034 Zea mays
21	894	55.7	1179	8	TOBANPER	W74105 Nicotiana s
22	887.5	55.3	1384	8	AF014469	AF014469 Oryza sat
23	877.5	54.7	1255	8	CCU12315	U12315 Genchirus ci
24	872	54.4	1850	8	AF149278	AF149278 Phaseolus
25	867.5	54.1	1220	8	AF014468	AF014468 Oryza sat
26	867.5	54.1	1220	8	AY085030	AY085030 Arabidops
27	849	52.9	978	8	AB024438	AB024438 Scutellar
28	848.5	52.9	1153	8	L1NPLXP	L24120 Linum usita
29	846.5	52.8	1184	8	SOPRKR5	Y10466 S.oleracea
30	844.5	52.6	1288	8	AF043235	AF043235 Striga as
31	838.5	52.3	1247	8	TAPERO	X56011 Wheat mRNA
32	835.5	52.0	149417	2	AP003991	AP003991 Oryza sat
33	834.5	52.0	1271	8	AF043234	AF043234 Striga as
34	831	51.8	43575	8	AP003817	AP003817 Oryza sat
35	822	51.2	930	8	HVU276227	AJ276227 Hordeum v
36	822	51.2	1354	6	AR208998	AR208998 Sequence
37	818	51.0	1285	8	HVPEROXI	X58396 Barley mRNA
38	816	50.9	91370	2	AP003757	AP003757 Oryza sat
39	812	50.6	1262	8	BLYPX	L36093 Barley pero
40	812	50.6	2388	8	AF014470	AF014470 Oryza sat
41	808	50.4	996	6	AX412623	AX412623 Sequence
42	808	50.4	1027	8	AY112238	AY112238 Arabidops
43	808	50.4	1126	8	AT286345	AT286345 Arabidops
44	808	50.4	1145	8	AF452384	AF452384 Arabidops
45	808	50.4	1231	8	AY074296	AY074296 Arabidops

## ALIGNMENTS

RESULT 1

LOCUS	CCU12314	1335 bp	mRNA	linear	PLN 02-AUG-1994
DEFINITION	Cenchrus ciliaris clone px7 peroxidase mRNA, complete cds.				
ACCESSION	U12314				
VERSION	U12314.1	GI:520567			
KEYWORDS					
SOURCE	Cenchrus ciliaris.				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade; Panicoideae; Paniceae; Cenchrus.				
REFERENCE	1 (bases 1 to 1335)				
AUTHORS	Ross, A.H.				
TITLE	Investigation of peroxidase genes and genetic transformation in buffel grass				
JOURNAL	Thesis (1994)	Botany, University of Queensland			
REFERENCE	2 (bases 1 to 1335)				
AUTHORS	Ross, A.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-JUL-1994)	Annette H. Ross, University of Queensland, Botany, St. Lucia, Brisbane, Queensland, 4072, Australia			
FEATURES	Location/Qualifiers				
source	1..1335				
	/organism="Cenchrus ciliaris"				
	/cultivar="Biloela"				
	/db_xref="taxon:35872"				
	/clone="px7"				
	/tissue_type="wounded and unwounded leaves and stems"				
	/clone_lib="lambda gt-10"				
	1..57				
5'UTR	58..981				
CDS	/EC_number="1.11.1.7"				
	/note="expressed in leaves and is wound-induced"				
	/codon_start=1				
	/product="peroxidase"				
	/protein_id="AA20477.1"				
	/db_xref="GI:520568"				
	/translation="MAAPTLMOCLLAISLSFTAAHSLSTFEVASGPNLOTIVRAMKVAVASSEPRMKSATLRLFFHDCFTYGGDCASTLNSGGKSAQNPANSVRYGEVLDITKKNVAACGGVASCATIVYPLARPGLNPGITWNPUGRUSTTAMSLANQNPPTKSLGTLISLIFGRLSARDMIALSGAHNAQRCTFRGRITGYDNTIDAFALQDQTCRPSGGGNLAPIDAOTPARPDNAYVNLNLSRGLFHSDELFNGSGQDALVROYSSPSQSFNSDVAAMIKMGINAGAGVRRRCRVNS"				
	58..126				
sig_peptide	/note="putative"				
	127..978				
mat_peptide	/product="peroxidase"				
	982..1335				
3'UTR	1335				
polyA_site	/note="39 A residues"				
	264 a 440 c 374 g 257 t				
BASE COUNT					
ORIGIN					
Alignment Scores:					
Pred. No.:	2,79e-86	Length:	1335		
Score:	1292.00	Matches:	255		
Percent Similarity:	89.14%	Conservative:	24		
Best Local Similarity:	81.47%	Mismatches:	28		
Query Match:	80.55%	Indels:	6		
DB:	8	Gaps:	4		
US-10-047-825-4 (1-313) x CCU12314 (1-1335)					
OY	1 MetAlaSerProThrIleuMetGInGlySerIleuValAlaIleSerIleuLeuSerCysValAla 20				
DB	58 ATGGCGGCTCTTACCTTGATGCAATCTCTGCTGGCCATCTCTCCCTCTCTTACCGCC 117				
OY	21 HisAlaGlnLeuSerProThrPheTyrAlaSerSerCysProAsnLeuGlnSerIleVal 40				
DB	118 CACGCCACACTGTGCACGACGCTTCTACGGGTCTCTCTGCCACACTGCAGACAGTCTT 177				
OY	41 ArgAlaIleMetThrGlnAlaValAlaSerGluGlnArgMetGlyAlaSerIleuArg 60				

Db	178	CGGGCCCGCCAGACCCAGAGGGCCGTGACAGAGCGAGCAAGCAAGATGGGGGCGCTTCATCCTCAGG	237
Qy	61	LeuPhepHeHISAspCySPhveAlGInGlySaSpGlySerIleLeuLeuAlaSPAlGly	80
Db	238	CTCTTTCTCCAGCAGCTGCTTGTGTCAGAGTGTGTAGCGCTTGATCTTTCTTCGACGCGGA	297
Qy	81	GlyGlyuThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlyValIleAspThr	100
Db	298	GGCGAGAGTCCGGCCGGCCGACCGCCAACTCGGTGGCGGGCTATGAGTATGACAGCC	357
Qy	101	IleuysArgAsnValGluAlaAlaCysProGlyValValIserCysAlaAspIleuAla	120
Db	358	ATCAAGAGAGACGTCGAGAGCCGCGCGCCAGCGCTGCTGTGGCGACCATGATGTCGCC	417
Qy	121	LeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGlyProThrTrpSerValIleGly	140
Db	418	CTCGCGGGGCGACCGCGGACCCGATTTCTTCGGCGGGCCGCACTGGAAAGCTGCCGTGGC	477
Qy	141	ArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThr	160
Db	478	CGCGCGGACGTCGAGACCGCGCGTGTCTGCTCCGCGGAACCGAACTCC3CGCGCGAGC	537
Qy	161	AlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGInGlyLeuSerProArgAspMet	180
Db	538	---ACCCGTGGGAGCCCTATCTCCTCTGTGGCGGACAG---CTTCCGCGCGGAGACATG	591
Qy	181	ThrAlaLeuSerGlyAlaAlaHisThrIleGlyAlaAlaArgCysThrThrPheArgGlyArg	200
Db	592	ATCGCGGCTGTGGCGGACACCAT---GCGCAGCGCCGATGACAGACAGCTCCGCGCGCGC	648
Qy	201	IleuYrGlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGInGlyThrCysPro	220
Db	649	ATCTACGGCGAGACCAACATGAGCCCTCTTGGCGGCGCTCCAGCAGAGAGCGTGCCG	708
Qy	221	ArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGInThrProValArgPheAsp	240
Db	709	CGGTGGCGCGGAGGAGGCAACTGGCGGCCCATCGAAGCGGAGACCGCGGAGGTTGCAC	768
Qy	241	ThrAlaYrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAspGInGlyLeu	260
Db	769	AACGCCCTACTACACACACTCTGTGCGGCGGGGCTTTTCTCACTCGGACGACGAGAGCTC	828
Qy	261	PheAsnGlyGlySerGInAspAlaLeuValArgGInYrSerAlaSerAlaSerLeuPhe	280
Db	829	TTCAACGGCGGGTCCAGAGACCGCGCTGTGTGAGAGCAGTACAGACAGACGCCCGTCGAGTTC	888
Qy	281	AsnAlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThr	300
Db	889	AACACGCACTTGTGTGGCGCCATGTATCAAGATGAGGCAACATATGGG-----GCCAAC	939
Qy	301	AlaGlyGInIleArgArgAsnCysArgValValAsnSer	313
Db	940	GCCGGACAGGTAGGCGCAACTGCGGGGTGTCAACAGC	978
RESULT 2			
LOCUS	AF155124	1369 bp	mRNA linear
DEFINITION	Gossypium hirsutum bacterial-induced peroxidase mRNA, complete cds		
ACCESSION	AF155124		
VERSION	AF155124.1	GI:5453378	
KEYWORDS			
SOURCE	Gossypium hirsutum.		
ORGANISM	Gossypium hirsutum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.		
AUTHORS	Assylbetse, K., Cuny, G., Valette, C., Delannoy, E., Bresson, E., Jallou, A., Daniel, J.-F., Geiger, J.-M. and Nicole, M.		
TITLE	Cloning and Characterization of a Bacterial-Induced Peroxidase-Encoding cDNA from Cotton (Accession No. AF155124).		
JOURNAL	Plant Physiol.	121 (1),	312 (1999)







	/organism="Oryza sativa (japonica cultivar-group)"
	/cultivar="Nipponbare"
	/db_xref="taxon:39947"
	/chromosome="6"
	/clone="OSUNBB0016D02"
BASE COUNT	44183 a 35607 c 35786 g 42641 t 239 others
ORIGIN	
Alignment Scores:	
Pred. No.:	4,37e-65 Length: 158456
Score:	1036.50 Matches: 218
Percent Similarity:	60.93% Conservative: 30
Best Local Similarity:	53.56% Mismatches: 56
Query Match:	64.62% Indels: 103
DB:	Gaps: 4
US-10-047-825-4 (1-313) x AP004731 (1-158456)	
OY 9 CysleuValAlaValserIleuLeuSerCysValAlaHisLa-----	22
Dd 76307 TGTAGTGTGCTGCTCTCTCTGCTACTTGTTGTGCAGATGGCCGAGTGTGATCACCGG 76366	
OY 22 -----	22
Dd 76367 CGGAGCCTTGATGGGCCCGCTGCTGTCGTGGTCCCGCTTTTGCTGGCGAGCGCGAG 76426	
OY 23 -----GlnLeuSerProThrPheTyrrAlaSerCysProAsnIeuInserIleVal 40	
Dd 76427 GCCGACAGGTGTCCCACTTCTACCTGAGACGTGCCCCAACCTGGCCACCATCTGTG 76486	
OY 41 ArgAlaAlaMetThrGlnAlaValAlaSerGluGlnArgMetGlyAlaSerIleuLeuArg 60	
Dd 76487 CGGTCCGGGTGGCTGCCGCCCTGGCGGAGGAGCCCGATGGGGGCTTCACCTTCCG 76546	
OY 61 LeuPhePheHisAspCysPheVal-----	68
Dd 76547 CTCTCTCTCATATTGCTCTCTCTCATATTAAAGACGACAACATTATTTTAATTAA 76606	
OY 68 -----	68
Dd 76607 TTTTTTGGCTGCTGTTGCTGTGAATGATCATGAGCTCATGAGTTATGGTTTATG 76666	
OY 69 -----GlnGlyAspCysIleuLeuAsp-----	78
Dd 76667 GTTGTCTGAGGGGCTTGTTGCGAGGATGTACAGGCTCGATCCGCTTGACGACAGCTG 76726	
OY 79 ---AlagIyGlyIuLystrThrAlaGlyProAsnIeuAsnSerValArgGlyPheGlyVal 97	
Dd 76727 ACGTTACCGCGGACAGAAGAGCCCGCCGCGAAGCCCACTGTGGCCGCGGGGTTCAGAGG 76786	
OY 98 IleAspThrIleLysTrgAsnValGlnAlaIlaAcysProGlyValIleSerCysAlaAsp 117	
Dd 76787 ATCGACGCCATCAAGACGCGAGGTGAGGCTCTCGAAGGCCACCGCTCTCCGCCGAC 76846	
OY 118 IleLeuAlaIleuAlaIlaArgSpsIlythrAsneu-----	129
Dd 76847 ATCTCGCCCTCGCCGCCGCGACGCGCTCACTGTCAGTACTACTACCTACCTGCTC 76906	
OY 129 -----	129
Dd 76907 TGACACGTACGTAAACCACCCGTACGTACATGAAATTTTACCAGTGGCTGATCTTCT 76966	
OY 130 -----LeuGlyIleProThrIrrPseerValIProLeuGlyArgArgAspSerThr 145	
Dd 76967 TTTCGTGTGACGTGGGTGGCCCACTGGAGCTGTGGCGCTGGGGCGGAAAGACTGGCCC 77026	
OY 146 ThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThrAlaSerIeuGlyThr 165	
Dd 77027 ACCGCGACGACGCGCGCGGCGAAGACGAACTCGCGGGGCGCCGGGTCTGCACCTCGCACG 77086	
OY 166 LeuIleSerIeuPheIlyArgGlnIlyIleuSerProArgSpsMetThrAlaIleuSerGly 185	
Dd 77087 CTCATCAGCATGTTCGGCACACAGGSCCTCTCGCGCGCGCATGACGAGCGCTGTGGGC 77146	

Oy	186	AlaHisThrIleGlyValIleArgCysThrRhpPheArgGlyArgIleTyrGlyAspThr	205
Db	77147	GCCACACCATTGGCGGCCGCCAGTCGCAGTTCTTCGCCAGCCGCATCTAACCGAGGCC	77206
Oy	206	AspIleAsnAlaSerPheAlaAlaLeuArgInGlnThrCysProArgSerGlyIAsp	225
Db	77207	AACATTCACGCCCTCTCCCTCCGTCCGCGAGCAAGACGTCCCGCGCTCCGGCGCGAC	77266
Oy	226	GlyAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyrPheThr	245
Db	77267	GCCAACTCCGCCGCTTGACGTGACAGAGCCGCCAGCGCTTCGACAACGCCCTACTACAG	77326
Oy	246	AsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnGluLeuPheAsnGlyGlySer	265
Db	77327	AACCTCTGTGCGAGCGCGCGCTGTCTCACATCCAGCACAGGAGGCTCTTAATGTGGCCTCG	77386
Oy	266	GlnAspAlaLeuValArgGlnIntYrSerAlaSerAlaSerLeuPheAsnAlaAspPheVal	285
Db	77387	CAGAGCGCGCTCGTAGCAGCATACAGCACCAACCCCTCCAGTCTCTCCGACTTCGCTGTC	77446
Oy	286	AlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArg	305
Db	77447	TCCCCCATTGGTCAAGTGGGCAACCTGCTCCGCTCCGACACCGCAGCGAGGTCAAG	77506
Oy	306	ArgAsnCysArgValAlaAsn	312
Db	77507	CTCAACTGCAGNAGGTGAAC	77527
RESULT 6			
SOPRX6			
LOCUS		1091 bp	mRNA
DEFINITION		S.oLERACEA mRNA for peroxidase, clone FC23.	linear
ACCESSION			PLN 15-JAN-1997
VERSION		Y10467	
KEYWORDS		Y10467.1 GI:1781331	
SOURCE		peroxidase; prxr6 gene.	
ORGANISM		Splinacla oleracea	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		Caryophyllidae; Caryophyllales; Chenopodiaceae; Splinacla.	
		1 (bases 1 to 1091)	
REFERENCE			
AUTHORS		Simon, P.	
TITLE		Diversity and conservation of plant peroxidases	
JOURNAL		Plant Peroxidase Newsletter 1, 4-7 (1993)	
REFERENCE		2 (bases 1 to 1091)	
AUTHORS		Simon, P.	
TITLE		Direct Submission	
FEATURES			
source			
		CH-1211 Geneva 4, SWITZERLAND	
		Location/Qualifiers	
		1..1091	
		/organism="Splinacla oleracea"	
		/sub_species="Model"	
		/db_xref="taxon:3562"	
		/clone="PC23"	
		/tissue_type="roots"	
		/tissue_type="leaves"	
		/tissue_type="stems"	
		/tissue_type="cotyledons"	
		/clone_lib="lambda UNIZAP-XR/SOCI"	
		/dev_stage="four-weeks-old plants grown in short days"	
		1..930	
		/gene="prxr6"	
		<1..930	
		/gene="prxr6"	
		/BC_number="1..11.1.7"	
		/codon_start=1	
		/evidence="experimental"	
		/product="peroxidase"	
		/protein_id="CAA71493.1"	
		/db_xref="GI:1781332"	





```

Db      4  GCATGTCACCTGGCTTCAGGAGCGCTCAGCGCCACTTTCAGCCACC 63
      32  SerCysProAsnLeuGlnSerIleValAlaGlnAlaMetThrGlnAlaValAlaSerGlu 51
      64  TCATGCCCAACCCCTCTCTACTATGTCGCCAATGCCATGACCAAGCGGTTAAATAGGAA 123
      52  GlnArMetGlyAlaSerLeuLeuArGLeuPhePheHisAspCysPheValGlnGlyCys 71
      124  AACCGCATGGCGCCCTCCATCCCGCTCCACTTTCATGACTGCTGCTTAATGGTTGT 183
      72  AspGlySerIleLeuLeuAsp-----AlaGlyGlyGlyValAlaGlyPro 87
      184  GATGGGTCGCTACTGTGGACGATGCTGCTCCTGCTGCTGCGAGAACAGCGAGGCC 243
      88  AsnLeuAsnSerValArGlyPheGlnValIleAspThrIleValArgAsnValAla 107
      244  AACCAAAAGCTCTGTAGAGGGTTCGACATATAGACAGATCAAGACAGAGTTGAGGCA 303
      108  AlacGlyProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThr 127
      304  GCATGCAACCCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
      128  AsnLeuGlnGlyIleProThrTrpSerValProLeuGlyArgAspSerThrThAla 147
      364  GTATTGGTAGCGGGCCACATGAGCGTCCACTAGGCCGAGCGGAGCGGCGGCGGCGG 423
      148  SerAlaSerLeuAlaAsnSerAsnProProProProProProProProProProPro 167
      424  AGCAGAGGCGCGCCAGCAACATCCAGCCGAGTCAAGCTAAGCAATGATACCATAC 483
      168  SerLeuPheGlyArGlnGlyLeuSerProArGAspMetThrAlaLeuSerGlyAlaHis 187
      484  AACTGTTCACGAACAAGGCGCTCACCGCTGACGTCACAAATCCCTGCGGGGCCAC 543
      188  ThrIleGlyAlaAlaArgCysThrThrPheAlaGlyArgIleGlyArgSerThrAspIle 207
      544  ACGATGCGGCGCGCGCTGACACAGCGTCCCGCAGCGATCTACAAAGACCAACATC 603
      208  AsnAlaSerPheAlaAlaLeuArGlnGlnThrCysProArGSerGlyAlaAspGlyAsn 227
      604  GACCGCGGCTGTCGCCACACCGCGAGGGGAGTCCCGGAGCGGAGCGGAGCGGAC 663
      228  LeuAlaProIleAspValGlnThrProValArGLeuAspThrAlaIleThrAsnLeu 247
      664  CTGCGACCGCTGAC---GGAACCCCGACCGCTGCGACCAACAGGTATACAGATCTG 720
      248  LeuSerArGArgGlyLeuPheHisSerAspGlnIleLeuPheAsnGlyGlySerGlnAsp 267
      721  GTGGCGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
      268  AlaLeuValArGlnIleThrSerAlaSerAlaSerLeuPheAsnAlaAspPheValAla 287
      781  GCCGTGGTGAAGCTTACAGCAACATGCGCTACTTTTGTACGAGATTTCGCCGCGCT 840
      288  MetIleArMetGlyAsnValGlyAlaLeuThrGlyThrAlaGlnIleArgAsn 307
      841  ATGGTAGAGTAGGCAACATCAGCCGTTGACCGGAGCAACGGGAGATCAGATTAAAC 900
      308  CysArGValValAsn 312
      901  TGCAGGAGGCGCTAAT 915

```

```

RESULT 8
LOCUS   AY065270      1238 bp      mRNA      linear      PLN 26-APR-2002
DEFINITION Arabidopsis thaliana putative peroxidase (At5g05340) mRNA, complete
cDS.
ACCESSION AY065270
VERSION   AY065270.1 GI:17529071
KEYWORDS  FLI, CDNA.
SOURCE    Arabidopsis thaliana.
ORGANISM  Arabidopsis thaliana

```

```

3
4

```

```

REFERENCE
AUTHORS
1 (bases 1 to 1238)
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banth,J., Chung,M.K.,
Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
Satou,M., Seki,M., Shinozaki,K., Southwick,A., Shinozaki,K.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 1238)
Yamada,K., Banth,J., Chang,C.H., Chang,E., Dale,J.M.,
Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinozaki,K., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

TITLE
JOURNAL
COMMENT

FEATURES
source
1..1238
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAF109-37-N16 (R18638)"
/notes="this clone is in a modified pluscript vector
(FUC-1) as a BamHI/XhoI insert.
ecotype: Columbia"
1..1238
/gene="At5g05340"
1..56
/gene="At5g05340"
57..1031
/gene="At5g05340"
/gene="At5g05340"
/codon_start=1
/evidence="experimental"
/product="putative peroxidase"
/protein_id="AA138746.1"
/db_xref="GI:17529072"
/translation="MASSKSLISILVLTLLDGGNNVVEAQLTFNFTSCPNLLS
TYQTAIVKSAVNSEARMGASILRIPEHDFVNGCGSILLDDTSFGEONAAANRNSA
RGENVINDTKSVEKACGCVSCADITAIARDVVALGCPNNVRYKRGADATASQA
AANSNIPAPFTSLSQLISSFSAVGLSTHDVVALGSAHTIGQSRNTRAKRTIYETIN
AAVATRTQRCPRASGSDGNLAPLDVTPAASFPNNFKNLMTQRLGHLSDVLFNGG
STDSIVRGYSNNPSSFNSDFTAAIMKMDISLPTGSSGETIRKVCGRTN"
1032..1238
/gene="At5g05340"
/misc_difference 1221
/gene="At5g05340"

```



Query Match: 60.63% Indels: 11  
 DB: 8 Gaps: 4  
 US-10-047-825-4 (1-313) x SOPRMR4 (1-1176)

QY 1 MetalaserProthreumMetGlnCysLeuValAlaValSerLeuLeuSerCysValAla 20  
 ||||| : : : : :  
 DB 38 ATGGCTTCGAACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 97

QY 21 HislaeIn-----LeuSerProthrePhetYrAlaSerCysProAnleu 36  
 ||||| : : : : :  
 DB 98 GGRACCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 157

QY 37 GlnSerIleValArgAlaAlaMetThrGlnAlaValAlaSerGluGlnArgMetGlyAla 56  
 : : : : :  
 DB 158 GAACAAATGTTTAAATAGCATGAGCAAGCGCTTAGTAGAGCAACGGGTAGGTGCT 217

QY 57 SerLeuLeuArgLeuPhePheHisaspCysPheValGlnGlyCysAspGlySerIleLeu 76  
 ||||| : : : : :  
 DB 218 TCCATACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 277

QY 77 LeuAsp-----AlaGlyGlyGlyValThrAlaGlyProAnleu---AsnSer 91  
 ||||| : : : : :  
 DB 278 TTGGATGATACACCACTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 337

QY 92 ValArgGlyPheGluValIleAspThrIleLysArgAsnValGluAlaIaCysProGly 111  
 : : : : :  
 DB 338 ATAAAGGGGTTTGGAGGATGATGCAATTAAGTTGAAGCTGCTGTTGCTGCA 397

QY 112 ValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuGly 131  
 ||||| : : : : :  
 DB 398 ACTGATGATGCGGAGCATCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 457

QY 132 GlyProThrTrpSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeu 151  
 ||||| : : : : :  
 DB 458 GGGCCCATGAGATGAAACTAGATAGAGAGATGCAAGAGACAAACATGACTTTA 517

QY 152 AlaAsnSerAsnProProProProProThrAlaSerLeuGlyThrLeuIleSerLeuPheGly 171  
 ||||| : : : : :  
 DB 518 GCAACCTAAATCTACACAGCAGGAGGAGTGCCTCCGCGCATCTGACGAGCACTGTTGCA 577

QY 172 ArgGlnGlyLeuSerProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGln 191  
 ||||| : : : : :  
 DB 578 AGCAAAATTTAAACATAGGAGATGACGCCCTGTGCGGAGGCCACACCATGGGTTT 637

QY 192 AlaArgCysThrThrPheArgGlyArgIleLysGlyAspThrAspIleAsnAlaSerPhe 211  
 ||||| : : : : :  
 DB 638 GCTCGGTGACCAACTCCGAGACCAATCTACACAGCAACGACATCGACCTACTT 697

QY 212 AlaAlaLeuArgGlnGlnThrCysProArgSerGlyGlyAspGly-----AsnLeuAla 229  
 ||||| : : : : :  
 DB 698 GGTGCCACAGCAGGAGCTACCTCCCTAGACCTACGAGCACTGCAACTGGGCC 757

QY 230 ProIleAspValGlnThrProValArgPheAspThrAlaLysPheThrAsnLeuSer 249  
 ||||| : : : : :  
 DB 758 CCAATGATATTCAAACTCCCAACACATTTGACAAATGATATTAACAAAACCTGTGGCG 817

QY 250 ArgArgGlyLeuPheHisSerAspGlnGluLeuPheAsnGlyLysSerGlnAspAlaLeu 269  
 : : : : :  
 DB 818 AAACGAGGATTAATCACTCCGATCAAGAGCTTACACAGGGGGTCTCAGAGCTGCTG 877

QY 270 ValArgGlnTySerAlaSerAlaSerLeuPheAsnAlaAspPheValAlaAlaMetIle 289  
 ||||| : : : : :  
 DB 878 GTCAATGATATTCAAACTCCCAACACATTTGACAAATGATATTAACAAAACCTGTGGCG 937

QY 290 ArgMetGlyAsnValGlyValLeuThrGlnIleThrAlaGlyGlnIleArgArgAsnGly 309  
 ||||| : : : : :  
 DB 938 AGAATGGGAGCTTAACCTTACTGATGATGAGGAGATTCGAAACAACTGCTGCG 997

QY 310 ValValAsn 312  
 ||||| : : : : :  
 DB 998 GTATCAAT 1006

RESULT 10  
 AB027752  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 Nicotiana tabacum (cultivar:Samsun NN) TMV-infected leaf mRNA,  
 clone.lib:lambda gt10 clone:tpoxcl.

REFERENCE  
 AUTHORS  
 TITLE  
 Hiraoka,S., Ito,H., Matsui,H., Honma,M. and Ohnishi,Y.  
 cDNA sequences for two novel tobacco peroxidase isoenzymes  
 (Accession Nos. AB027752 and AB027753). (PGR99-109)  
 Plant Physiol. 120, 1205 (1999)

JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

FEATURES  
 source  
 1..1356  
 /organism="Nicotiana tabacum"  
 /cultivar="Samsun NN"  
 /db\_xref="taxon:4097"  
 /clone="tpoxcl"  
 /tissue\_type="TMV-infected leaf"  
 /clone.lib="lambda gt10"  
 45..1010  
 /EC\_number="1.11.1.7"  
 /codon\_start=1  
 /product="peroxidase"  
 /protein\_id="BA82306.1"  
 /db\_xref="GI:5381253"  
 /translation="MASLKINALVLFVLSLIGSSAOLSTGFSKSPKLYOTVKS  
 ANQAIKRPKMSALIRFHFQVNGCGSLLDDTSFEKRAAPVNSARPE  
 VIDIKASVKEVCGVYSCADITLAIVRDSVVLGGEPMNVKLRGRDSRTASANS  
 GIPRATSNLRLLISFSAVGLSTRDWALSGAHTTQARTSRANIVETNNLDSF  
 AKTRQSNCPRSSGSDNNLALPLDQTPNKEDNNYFKNLVDRKGLHSDOLFNGGSAD  
 SIYTSYNNSSSSSDPVTAMIKMGDIPKPLTSGNGEIRKRCRLN"

BASE COUNT  
 376 a 257 c 266 g 457 t

ORIGIN

Alignment Scores:  
 Pred. No.: 9.66e-62 Length: 1356  
 Score: 958.50 Matches: 190  
 Percent Similarity: 74.1% Conservative: 39  
 Best Local Similarity: 61.49% Mismatches: 73  
 Query Match: 59.76% Indels: 7  
 DB: 8 Gaps: 3

US-10-047-825-4 (1-313) x AB027752 (1-1356)

QY 11 ValAlaValSerLeuLeuSerCysValAlaHisIaGlnLeuSerProthrePhetYrAla 30  
 : : : : :  
 DB 81 ATTTAGTGAATTAATTAATGAAGCTCACTCACTCACTCACTCACTCACTCACTCA 140

QY 31 SerSerCysProAnleuGlnSerIleValArgAlaAlaMetThrGlnAlaValAlaSer 50  
 ||||| : : : : :  
 DB 141 AAATCTTGCTTAACCTTATCAACATGTAAGACAGAGCAATGCTATTAACAAG 200

QY 51 GluGlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHisaspCysPheValGlnGly 70  
 ||||| : : : : :  
 DB 201 GAACAGAAATGGGCGTCTCTCTGCTCATTTCCATGATGATGCTCGTCAATGGA 260

QY 71 CysAspGlySerIleLeuLeuAsp-----AlaGlyGlyGlyValThrAlaGly 86  
 ||||| : : : : :  
 DB 998 GTATCAAT 1006

Db 261 TGGATGATCGCTACTCTCATGATACATCAAGCTTCACAGAGAGAAAGGCTGCA 320  
Qy 87 ProbsnleuasnserValArgglyPhegluValIlaaspthrIlelysarasnValGlu 106  
Db 321 CCAAAATCTTAATCTGCTAGAGGATTGGAATTAATGACAAATATCAAACTGCTGGAG 380  
Qy 107 AlalacysProglValValSerCysAlaasphleuAlaLeuAlaAlaArgaspGly 126  
Db 381 AAGTGTGCCCTGCTGTGTTCTTCTGATATTTTGGCCCTGCTGCGGACTCT 440  
Qy 127 ThrAsnleuLeuGlyGlyProthrTrpSerValProleuGlyArgargaspSerThrTrp 146  
Db 441 GTTGCTACTTGTGGAGAGTCCCATTTGGAATGTAATTTGGAGAGAGAGATTCTAGACT 500  
Qy 147 AlaSerAlaSerLeuAlaasnsersnProProProProThrAlaSerleuGlyThrLeu 166  
Db 501 GCAAGTCAAGGTCTGCGCAATAGTGCAATTCCTGCTGACTCTTAACCTTATAGACGC 560  
Qy 167 IleSerleuPheglValArgglyLeuSerProArgaspMetThrAlaLeuSerGlyAla 186  
Db 561 ATCTCAAGTTTCAGTCTGCTGCTGCTTTCACCAAGAGATATGCTTATCTGGGCT 620  
Qy 187 HisThrIleGlyAlaAlaArgCysThrThrPheArgGlyArgIleThrGlyAspThr--- 205  
Db 621 CACACAAATTGACACAGCAAGATGATACATCTTTAGGGCACGTRATACACAGACCAAT 680  
Qy 206 AspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSer-----Gly 223  
Db 681 AACTAGATGCATCATTTGACAGCACTAGCAAAAGCAACTGCCCAAGAGATTGAGCTCA 740  
Qy 224 GlyAspGlyAsnleuAlaProIleasphValGlnThrProValArgPheAspThrAlaTrp 243  
Db 741 GGGGACACAAATTGGCACCTGATGTCACCAACCTCAACAAATTATACAACAATTTAT 800  
Qy 244 PheThrAsnleuLeuSerArgArgGlyLeuPheHisSerAspGlnleuLeuPheasnGly 263  
Db 801 TTCAGAAATCTTGTGACAAAAGAGGCTCTTCTGACTGATCAACAACTTTATATGT 860  
Qy 264 GlySerGlnAspAlaLeuValArgGlnThrSerAlaSerAlaSerleuPheAsnAlaAsp 283  
Db 861 GGATCTGCCCATTCGATGTTGACATCTTACATCAACATCTGAGTTCAGTCTGAT 920  
Qy 284 PheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGln 303  
Db 921 TTGTGACTGCTGATGATTAAGATGGTGATATGTCACCTTACTGTTCTAATGAGAG 980  
Qy 304 IleArgArgAsnCysArgValAlaAsn 312  
Db 981 ATTAGAGAAATTGCCGTAGGCTCAT 1007  
RESULT 11  
AB024439 957 bp mRNA linear PLN 07-OCT-1999  
LOCUS Scutellaria baicalensis mRNA for peroxidase 3, complete cds.  
DEFINITION AB024439  
ACCESSION AB024439.1 GI:4760703  
KEYWORDS peroxidase 3.  
SOURCE Scutellaria baicalensis cDNA to mRNA.  
ORGANISM Scutellaria baicalensis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Lamiales; Lamiaceae; Scutellaria.  
REFERENCE 1 (sites)  
AUTHORS Morimoto,S., Tateishi,N., Inuyama,M., Taura,F., Tanaka,H. and Shoyama,Y.  
TITLE Identification and molecular characterization of novel peroxidase  
JOURNAL J. Biol. Chem. 274 (37), 26192-26198 (1999)  
MEDLINE 99403061  
REFERENCE 2 (bases 1 to 957)  
AUTHORS Morimoto,S.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAR-1999) Satoshi Morimoto, Kyushu University,

Faculty of Pharmaceutical Sciences, Higashi-Ku Maidsashi 3-1-1,  
Fukuoka, Fukuoka 812-8582, Japan  
(E-mail: morimoto@shoyaku.phar.kyushu-u.ac.jp, Tel:81-92-642-6581,  
Fax:81-92-642-6545)  
FEATURES  
source location/Qualifiers  
1..957  
/organism="Scutellaria baicalensis"  
/db\_xref="taxon:65409"  
1..957  
/codon\_start=1  
/product="peroxidase 3"  
/protein\_id="BA077389.1"  
/db\_xref="GI:4760704"  
/translation="MASFVSEFSTRMLVLLIGVNAQUSANFYNTCPNLITRN  
AVNSAVSDTRMGASLDLRFHDCFVGCASVLLDRTGTEKTAQPNANLRFED  
VIDIKTIVGSGCNVYSCSDILSVARSDVAVGVSMAVLAGRDSSTASNAAT  
OIPBGLNVALITPSFNKFTAREWALVSGHRTIGARCTGRTYNDNTNGAPA  
TGLRANCPBSGGDNLPILNVSPPARNNDYIRNLIGLRGLSDDELNNGTADAOV  
RAYSTNSAAFFNDPANAMVMSNLSPLTNGOIRRCRRTN"  
BASE COUNT 237 a 252 c 245 g 223 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.72e-61 Length: 957  
Score: 950.00 Matches: 184  
Percent Similarity: 74.27% Conservative: 44  
Best Local Similarity: 59.93% Mismatches: 75  
Query Match: 59.23% Indels: 4  
DB: Gaps: 1  
US-10-047-825-4 (1-313) x AB024439 (1-957)  
Qy 10 LeuValAlaValSerleuLeuSerCysValAlaHisAlaGlnleuSerProThrPheTrp 29  
Db 34 TTAATGCTGCTTCTCTTTAATGAGAGTACGATGACAAATGCTGCAAAATTTCTAC 93  
Qy 30 AlaSerCysProAsnleuGlnSerIleValArgAlaAlaMetThrGlnAlaValAla 49  
Db 94 AACCAACATGCTCAATCTTCAATGATTTATGGAACGCTGTAACCTGCTGCTGCG 153  
Qy 50 SerGlnArgMetGlyAlaSerleuLeuArgPhePheHisAspCysPheValGln 69  
Db 154 ACTGATCTGATGAGGGGCTGCTGCTCCGCTCCATTTCCATGATGCTTGTGCTCAAT 213  
Qy 70 GlyCysAspGlySerIleLeuLeuAsp-----AlaGlyGlyGluGlyThrAla 85  
Db 214 GGGTCGATGATCGGCTTTACTAGATGACAGAGAGCTGGCTCACCGGAGAAAGAGGCC 273  
Qy 86 GlyProAsnleuAsnserValArgglyPhegluValIlaaspthrIlelysarasnVal 105  
Db 274 GGTCCCAATGCTAATTTCCGTCGAGGGGTTGATGCAATATACAGAACTAGTACTAGT 333  
Qy 106 GluAlaAlaCysProglValValSerCysAlaasphleuAlaLeuAlaAlaArgasp 125  
Db 334 GAGGGGTGTCGCCCAACATGCTCTCTGCTCCGATATTTCTTCCGTTCAGCCAGGAG 393  
Qy 126 GlyThrAsnleuLeuGlyGlyProthrTrpSerValProleuGlyArgargaspSerThr 145  
Db 394 GCGCGTTCGCTGTTGGGGGTCCAAAGTGGGACATGACCTAGACAGAGAGGACTCAAC 453  
Qy 146 ThrAlaSerAlaSerleuAlaasnsersnProProProProThrAlaSerleuGlyThr 165  
Db 454 ACAGCTAATCTGAATGACGACCAACCACTCCAGAGCTGAGCTCAATCTCAATAGCC 513  
Qy 166 LeuIleSerleuPheglValArgglyLeuSerProArgaspMetThrAlaLeuSerGly 185  
Db 514 CTCATCACTTCTTCTCAACAAAGCTTCACTGCGCAGAGAAAGTTGCTCTTTCGGA 573  
Qy 186 AlaHisThrIleGlyAlaAlaArgCysThrThrPheArgGlyArgIleThrGlyAspThr 205  
Db 574 TCCACACAGATAGGACAGGAGGTCAGACGATTCGTTGGAAGAGATCAACAGACAG 633  
Qy 206 AspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyAsp 225

```

Db 634 AACATTAAGGAGCGCTGCCGACGCGGCGTGAAGGCACTGCGCCGAGGAGGAGAC 693
      :::::::::: :::::::::: ::::::::::::::::::::::::::::::
Qy 226 GYAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyPheThr 245
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 694 AACACCTCGCCCGCTGGACAACTGAGCCCGGCGGCAACAGATTATACAGG 753
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 246 AsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnGluLeuPheAsnGlyLys 265
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 754 AACCTGATGGGTGTAAGGGGTCTGCTCCACTAGACAGGAGCTTTCAACATGGGACC 813
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 266 GlnAspAlaLeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsnAlaAspPheVal 285
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 814 GCGGACGCTCAGGTGACAGCTTACAGACCACTCGGCTGCTTTCACAGATTTCGCC 873
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 286 AlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArg 305
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 874 AACGCCATGGTGAAGATGTCACACTTACGCCACTGACGCCAAGTGTCAGATCAGG 933
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 306 ArgAsnCysArgValValAsn 312
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 934 AGGAAGCTGCAGGAGAACTAAC 954
      ::::::::::::::::::::::::::::::::::::::::::::::

RESULT 12
SSNPEROXIA 1144 bp mRNA linear PLN 01-AUG-1995
LOCUS SSNPEROXIA Stylosanthes humilis peroxidase mRNA.
DEFINITION L36110.1 GI:535762
ACCESSION L36110.1 GI:535762
VERSION L36110.1 GI:535762
KEYWORDS peroxidase.
SOURCE Stylosanthes humilis (strain Paterson) cDNA to mRNA.
ORGANISM Stylosanthes humilis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Stylosanthes.
REFERENCE 1 (bases 1 to 1144)
AUTHORS Harrison,S.J., Curtis,M.D., McIntyre,C.L., Maclean,D.J. and
Hanners,J.M.
TITLE Differential expression of peroxidase isogenes during the early
stages of infection of the tropical forage legume Stylosanthes
humilis by Colletotrichum gloeosporioides
JOURNAL Mol. Plant Microbe Interact. 8 (3), 398-406 (1995)
MEDLINE 95383712
PUBMED 7655062
FEATURES
source location/Qualifiers
1..1144
/organism="Stylosanthes humilis"
/strain="Paterson"
/db_xref="taxon:35628"
/clone="Shox6"
/tissue_type="stem"
1..1144
BASE COUNT 360 a 205 c 222 g 357 t
ORIGIN
mRNA
BASE COUNT 360 a 205 c 222 g 357 t
ORIGIN

Alignment Scores:
Pred. No.: 3.3e-60 Length: 1144
Score: 936.50 Matches: 187
Percent Similarity: 73.08% Conservative: 41
Best Local Similarity: 59.94% Mismatches: 75
Query Match: 58.39% Indels: 9
DB: 8 Gaps: 4

US-10-047-825-4 (1-313) x SSNPEROXIA (1-1144)
Qy 9 CysLeuVal-----AlaValSerLeuLeuSerCysValAlaHisAlaGlnLeuSerPro 26
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 69 TGGTTATATAATATGTGTGATGACCTTATAGGA---TTAGATCAGGTCATATGTCATCA 125
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 27 ThrPheTyrAlaSerSerCysProAsnLeuGlnSerIleValArgAlaAlaMetThrGln 46
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 126 AATTTTTATGCAACACATGTCGAAATGCATTTCAAGCATTTAGTACAGAGTGAATCT 185
      ::::::::::::::::::::::::::::::::::::::::::::::

```

```

Qy 47 AlaValAlaSerGlnGlnIleArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAspCys 66
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 186 GCTGTAGCAAAAGAGCTCGCATGGAGCTTCCCTTCTTGCTTCAATTCATTCATGATTC 245
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 67 PheValGlnGlyCysAspGlySerIleLeuLeuAsp-----AlaGlyGlyGlu 82
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 246 TTTGTTCAAGATGTGATGATCATGCTGTATTATGATGATACATCAATTTCCACAGAGAA 305
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 83 LysThrAlaGlyProAsnLeuAsnSerValArgGlyPheLeuValIleAspThrIleLys 102
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 306 AAAACACACAGCTCTAATGCTTAATTAAGGTTTAAAGTCATAGCACACCAATAAA 365
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 103 ArgAsnValGlnAlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAla 122
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 366 TCTCAAGTAGAGACTGTGTGCTGCTGCTGCTTCTGCTGCTATATCTTGCTGTTGGCT 425
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 123 AlaArgAspGlyThrAsnLeuLeuGlyGlyProThrTyrSerValProLeuGlyArgArg 142
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 426 GCTAGAGATTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 143 AspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThrAlaSer 162
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 486 GACTCACTACAGCAAGCTTAAAGCTTACCTACATTCGCTGCTGCTGCTGCTGCTGCTGAT 545
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 163 LeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThrAla 182
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 546 CTGACGTGCTCAATCTCTGCTTCTCTACAGAAAGTTATCACTAGTGAAGTTGGTCC 605
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 183 LeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArgGlyArgIleTyr 202
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 606 CTATCAGAGAGGCAATCAATATGGGCAAGCAAGATGCAAGCTTTAGAACACAGATATAC 665
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 203 GlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSer 222
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 666 ACTAGAGCAACATGATATCCCAATTTGGCCAAATCATGCAAGAAATTCCTCATATAC 725
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 223 -----GlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrProValArgPheAsp 240
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 726 ACAGCAATGTTGTCACACAACTTGGCCCAATTCACACAACTGTCACAAAGATTGGAC 785
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 241 ThrAlaTyrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnLeu 260
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 786 AATGTTACTATTAAGAACTGCTGTAGTAAAGAGGCTCTTCCTGCTGATCAACAACATC 845
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 261 PheAsnGlySerGlnAspAlaLeuValArgGlnTyrSerAlaSerAlaSerLeuPhe 280
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 846 TTCAATGAGGATCCACAGATTCTCAAGTGAATGTTATGCTTCAACCTTCAAGTTTC 905
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 281 AsnAlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThr 300
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 906 TGCCTGTATTGTCGCAATGCTATGATTAAGATGTTAACTATGATCCACACATGATCC 965
      ::::::::::::::::::::::::::::::::::::::::::::::
Qy 301 AlaGlyGlnIleArgArgAsnCysArgValValAsn 312
      ::::::::::::::::::::::::::::::::::::::::::::::
Db 966 AGTGGCAGATTAGACCAATTCGAGGAGACCAAT 1001
      ::::::::::::::::::::::::::::::::::::::::::::::

RESULT 13
AF247700 1330 bp mRNA linear PLN 03-JUL-2000
LOCUS AF247700 Oryza sativa peroxidase POCI mRNA, complete cds.
DEFINITION AF247700
ACCESSION AF247700
VERSION AF247700.2 GI:8901179
KEYWORDS
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 1330)
AUTHORS Hlaire,E., Young,S.A., Willard,L.H., McGeer,J.D., Sweat,T.A.,
Chittoor,J.M., Gulikema,J.A. and Leesch,J.E.
TITLE Induction of peroxidase gene POCI during the incompatible

```





/protein\_id="CAA46916.1"  
 /db\_xref="gi:20286"  
 /translation="MSATNSLSLMLTVAAMASVSAOLATFVPTSCGNALSTIK  
 SVITVAASSEARMGASILRIHPDQFQGCASVILSQEONAGPNVGLGFSVIND  
 AKAPVLAICNCTVSCADLLAVARDSVVALGSPVTLVGRDSTLSAALANTDPA  
 PSSSLAEILGNSTRKGLDAPDAITVLSGHTIGQACQCFRURITENINIDAEITQK  
 ANCPVPTSGSDSLIAPVDTTPNAPDNVYSNLLSNKGLHSDVLFNGSGADTVERN  
 FASNAAPFSAVTAMVMGNISPLTGTGQIRLSCSKVNS"  
 22..96  
 sig\_peptide  
 sig\_peptide  
 97..972

BASE COUNT	321 a	377 c	353 g	236 t
ORIGIN				

Alignment Scores:	
Pred. No.:	1,59e-59
Score:	928.00
Percent Similarity:	72.33%
Best Local Similarity:	59.43%
Query Match:	57.86%
DB:	8
Gaps:	3
Length:	1287
Matches:	186
Conservative:	44
Mismatches:	82
Indels:	6
Gaps:	3

US-10-047-825-4 (1-313) x OSPER (1-1287)

QY	1	MetaLsePrOthr-----LemMelncSueVuaLalaLseIreneuSer	17
Db	22	ATGGCTTCGCAACTAATTTCTTCACTTAAGCTAATGCTGCTGGTGGCTCACTAATGGC	81
QY	18	CysValAlaAlaAlaGlnLeuSerProthrPheTyrAlaSerSerCysProAlaLeuGln	37
Db	82	TCAGTGGCAATCGCCCACTCTCGGCAAGCTTTATGACACGTCGTCGCCCAATGCTTG	147
QY	38	SerTLeValArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnAArgMetGlnAlaSer	57
Db	142	TCACACATCAAAAGCGTATATACAGCCGCGCTGAACAGCAGAGCTGCCATGGGGCGCG	201
QY	58	LeuLeuArgLeuPhePheHisASPcysPheValGlnGlyCysAspGlySerTLeuLeu	77
Db	202	CTGCTAGGCTCACTCTTCCAGACGTGTTGCTCCAAAGGGGCGAGCGCCAGTTCTGCTG	261
QY	78	AspAlaGlyGlyLeuThrAlaGlyProAlaLeuAlaSerValArgGlyPheGlnVal	97
Db	262	----TCGGGCAAGACCAAAATGCTGGCCCAAGAGTGGGTGGTGAAGGATATCAAGCTC	318
QY	98	ILeAspThrTLeuArgAsnValGlnAlaAlaCysProGlyValValSerCysAlaAsp	117
Db	319	ATCGAACACGCCAAGCGCGGTGGAGGCCATCTGAACACAGACGTCCTCGCGCGGAC	378
QY	118	ILeLeuAlaLeuAlaAlaArgAspGlyTThrAsnLeuGlnGlyProThrTTPSerVal	137
Db	379	ATCTCGCGGTGGCGCCGCGAGCTCCGTCGCTGGAGGGCCATCATGACGGTT	438
QY	138	ProLeuGlyArgArgAspSerThrThAlaSerAlaSerLeuAlaAsnSerAsnPro	157
Db	439	CTGTTGGGAGAAAGGAAGTCCACACATCGCAAGGAGGCTTGCAAAATACGACACTCCT	498
QY	158	ProProThrAlaSerLeuGlyTThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerPro	177
Db	499	GCCCTCTCTTACGCTCGCAGAACTTATGGCAATTTTCCAGAAAGGAGCTGACGCA	558
QY	178	ArgAspMetThrAlaLeuSerGlyAlaAlaHisThrILeGlyGlnAlaArgCysThrThrPhe	197
Db	559	ACCGACATGGTTCCTCTCTAGAGACACACACATGTGGGCAAGGCGCAATGCCAATTTC	618
QY	198	ArgGlyArgGlyTyrGlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGln	217
Db	619	AGGAGACAGATCTCAACAGAACCAATCATGATTCGCTTGCGAGCCACAGCCAGGCC	678
QY	218	ThnGysProArg-----SerGlyGlyAspGlyAsnLeuAlaProTLeAspValGlnThr	235
Db	679	AACTGGCCACGCGGACGGGCGAGGAGACAGCAACCTTGGCGCCCTGGACAGCAGCAGC	738

OY	236		ProvalArgPhasphrIatrrPrhemrarranleuleSerAargArGlyLeupePheHts	255
Dd	739	CCCAACGCCCTTGACAACCCTTACAGCACAACCTGTCTCCAACAAAGGGCTCTGCAC		798
OY	256	SerAspGLInglLeuPheasnlglySerglAnspAlalaLeuValArgIntyrSerAla		275
Dd	799	TTCGAACCAAGTGCTCTTAACGGCGGCACGGGCAACACGGTCAAGAATCTTCGCTCC		858
OY	276	SerAlaSerLeuPheasnAlaspheValAlalMetileArgnecglysnValgily		295
		: : : : : : : :     : : : :             : : : : : : : : : : : : : : : : :		
Dd	859	AACGGCGCGCTTAGCAGCGCCCTTCAAGCACGGCCCATGTGTAAGATGGGGAACATCTCG		918
OY	296	ValLeuThrGlyThralalaglnllearGrarsnCyrsArgValValnsSer		313
		: : :		
Dd	919	CCGCTGACCGGAGCAGGCGCAGATCAAGCTCAAGTCAAGSTGAACITCC		972
RESULT 15				
LOCUS	ARCPNC1	1116 bp	mRNA	linear PLN 15-AUG-1996
DEFINITION	Archis hypogaea cationic peroxidase (PNC1) mRNA, complete cds.			
ACCESSION	M37636			
VERSION	M37636.1 GI:1491775			
KEYWORDS	cationic peroxidase.			
SOURCE	Archis hypogaea cDNA to mRNA.			
ORGANISM	Archis hypogaea			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	Rosidae; eurosids I; Fabales; Fabaaceae; Papilionoideae;			
	Aeschynomeneae; Arachis.			
	1 (bases 1 to 1116)			
	Buffard,D., Breda,C., van Huyssee,R.B., Asemota,O., Pierre,M.,			
	Ha,D.B. and Esnault,R.			
TITLE	Molecular cloning of complementary DNAs encoding two cationic			
JOURNAL	peroxidases from cultivated peanut cells			
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 87 (22), 8874-8878 (1990)			
PUBMED	91062381			
REFERENCE	2 (bases 1 to 1116)			
AUTHORS	Esnault,R.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-AUG-1990) Robert Esnault, Institut des Sciences			
	Vegetales, CNRS, 91198 GIF sur Yvette, France			
	On Aug 15, 1996 this sequence version replaced gi:166472.			
COMMENT	Location/Qualifiers			
FEATURES	1..1116			
source	/organism="Arachis hypogaea"			
	/db_xref="taxon:3818"			
	/clone="prxpnc1"			
	1..1116			
gene	/gene="PNC1"			
mRNA	<1..1116			
CDS	/gene="PNC1"			
	13..963			
	/gene="PNC1"			
	/codon_start=1			
	/product="cationic peroxidase"			
	/protein_id="AA806183.1"			
	/db_xref="GI:1491776"			
	/translation="MALPIKRVDFLIWCLIGLSAQSSNFYATKPALSTIKSAV			
	NSAAEARMGASLRLRHFDHFDCFOQCDASVLDDTSNFGERTAGPNANSIRFEVIY			
	DTISGVESLCPGVASCADILIAVARSDVALGGASWNVLGRDSTLASLNSDL			
	PAPFENLGLISAFAFSNGKGFTEKLIVLTSGAHHTGOACTAFPRITYSNIDPYAKS			
	LQANCPSEVGDJNISPDPTTPKNFNDAAYYNLNPKKGLLHSQQOLENGYSTDSQVITA			
	YSNNAAFNTDFGNAMTKMGLSLPTSITSQIRINKRKTN-			
BASE COUNT	349 a 214 c 213 g 340 t			
ORIGIN				
Alignment Scores:				
Fred. No.:	1.6e-59	Length:	1116	
Score:	927.00	Matches:	182	
Percent Similarity:	71.70%	Conservative:	41	
Best Local Similarity:	58.52%	Mismatches:	78	



Query Match: 57.79% Indels: 10  
DB: 8 Gaps: 2

US-10-047-825-4 (1-313) x ARCPN1 (1-1116)

```
QY 6 LeuMetGlnCysLeuValAlaValSerLeuSerCysValAlaHisAlaGlnLeuSer 25
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 46 ATATTATATGCTCTTATAGGA-----TTAGGGTCAGCTCATTTGCTCA 87
QY 26 ProThrPheThrAlaSerSerCysProAsnLeuGlnSerIleValAlaAlaMetThr 45
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 88 TCTAATTTTATGCACAAATGTCCTCAATGCATTTCACAACTTAAGTCACAGTGAAC 147
QY 46 GlnAlaValAlaSerGlnGlnArgMetGlyAlaSerLeuLeuArgLeuPheHisAsp 65
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 148 TCTGCTGCTGCCTCAAGAGAGCTGAGAGCTCCCTTCTTGGCTTCATTTCATGAT 207
QY 66 CysPheValGlnGlyCysAspGlySerIleLeuLeuAsp-----AlaGlyGly 81
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 208 TGCTTTGTTCAAGATGTGATGTCATGATGATGATGATGATGATGATGATGATGATGAT 267
QY 82 GlnGlyThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlnValIleAspThrIle 101
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 268 GAAAGACACAGCAGAGTCCAATGCAATTCATTAAGAGAGTTTGAAGTATGATACATA 327
QY 102 LysArgAsnValGlnAlaAlaCysProGlyValValSerCysPalaAspIleLeuAlaLeu 121
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 328 AAGTCCCAAGTAGAGAGCTTGCTGCTGAGAGAGCAAGTTGGAATGTGTTATGCTGTT 387
QY 122 AlaAlaArgAspGlyThrAsnLeuGlnGlyProThrTrpSerValProLeuGlyArg 141
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 388 GCTGCTAGAGACTCTGTGTTGCTAGAGAGAGCAAGTTGGAATGTGTTATGCTGTT 447
QY 142 ArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThrAla 161
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 448 AGAGACTCAACACACTGCAATTAAGCTGCTGACTGATTTGCCGCTCCATTGTTT 507
QY 162 SerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThr 181
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 508 AATCTTAGTGGCTTATCTGTCTTCTCCACAAAGGTTTCACAAACAAAGACTGTT 567
QY 182 AlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArgGlyArgIle 201
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 568 ACCTATACAGAGCGGCTATCAATTTGGGCAAGCAACATGACAGCTTCAGAACAGAGATT 627
QY 202 TyrGlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArg 221
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 628 TACAAATGAGAGCAACATGATCCAAATATGCAAAATCATTCAGAACCAAAATGCTCTAGC 687
QY 222 SerGlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrProValArgPheAspThr 241
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 688 GTAGAGAGAGATACCAATTTGTCCACCTTTGATGATACAAACACCAACAAATTTGACAAAT 747
QY 242 AlaTyrPheThrAsnLeuSerArgArgGlyLeuPheHisSerAspGlnGlnLeuPhe 261
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 748 GCTTACTATATCAACTGAGAACAGAGGCTCTTGATCTGATCAACAACTCTTC 807
QY 262 AsnGlyGlySerGlnAspAlaLeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsn 281
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 808 AATGCTGATATCCATGCTCCAGTCACTGCTTATAGCAACATGCTGCAACTTTCAC 867
QY 282 AlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAla 301
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 868 ACTGACTTTGCAATGCAATGATTAAGATGGAACCTTAGTCACACTGAGGAGCTAGT 927
QY 302 GlyGlnIleArgArgAsnCysArgValValAsn 312
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 928 GGCCAATTTAGAACCAATTCAGAGAACCAAC 960
```

Search completed: May 3, 2003, 12:58:50  
Job time : 2264 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

SUMMARIES

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 3, 2003, 11:27:38 ; Search time 190 Seconds

(without alignments)  
3709.871 Million cell updates/sec

Title: us-10-047-825-4

Perfect score: 1604

Sequence: 1 MASPTLMQCLIAVAVSLSCVA.....VGVLGTAGIRNRCHVNS 313

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/uspro.spool/US10047825/rnatc.26042003.111957.7124/app.query.fasta.1.455  
-DB=N\_Geneseq.101002 -OPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPT=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10047825.0CGN.1.1.200.0.rnatc.26042003.111957.7124 -MCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEOUTERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq.101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Score Match Length DB ID Description

1	982.5	61.3	1156	22	AAH44077	Oryza sativa perox
2	967	60.3	1224	22	AAH44087	Pinus radiata pero
3	954.5	59.5	1158	22	AAH44080	Oryza sativa perox
4	950	59.2	1144	18	AAH44085	Stylosanthes humil
5	936.5	58.4	1171	22	AAH44082	Oryza sativa perox
6	926	57.7	1310	22	AAH44085	Oryza sativa perox
7	901	56.2	1171	22	AAH44082	Oryza sativa perox
8	899.5	56.1	1391	21	AAH44081	Pinus radiata pero
9	896.5	55.9	1379	22	AAH44083	Pinus radiata pero
10	885.5	55.2	1223	22	AAH44083	Pinus radiata pero
11	870.5	54.3	1218	21	AAH44083	Pinus radiata pero
12	867.5	54.1	1220	21	AAH44083	Pinus radiata pero
13	846	52.7	801	21	AAH44083	Pinus radiata pero
14	822	51.2	1355	20	AAH44083	Pinus radiata pero
15	818	51.0	1295	24	AAH44083	Pinus radiata pero
16	813	50.7	1230	22	AAH44081	Pinus radiata pero
17	807	50.3	916	21	AAH44081	Pinus radiata pero
18	801	49.9	1244	21	AAH44081	Pinus radiata pero
19	801	49.9	1247	21	AAH44081	Pinus radiata pero
20	778	48.5	1186	21	AAH44081	Pinus radiata pero
21	774	48.3	1014	21	AAH44081	Pinus radiata pero
22	771.5	48.1	1294	21	AAH44081	Pinus radiata pero
23	755.5	47.1	1351	21	AAH44081	Pinus radiata pero
24	739.5	46.1	1073	21	AAH44081	Pinus radiata pero
25	727.5	45.4	1244	19	AAH44081	Pinus radiata pero
26	727.5	45.4	1244	19	AAH44081	Pinus radiata pero
27	725	45.2	1183	21	AAH44081	Pinus radiata pero
28	723	45.1	975	22	AAH44081	Pinus radiata pero
29	721.5	45.0	1270	16	AAH44081	Pinus radiata pero
30	721.5	45.0	1270	16	AAH44081	Pinus radiata pero
31	721.5	45.0	1270	16	AAH44081	Pinus radiata pero
32	719.5	44.9	911	21	AAH44081	Pinus radiata pero
33	717	44.7	1348	14	AAH44081	Pinus radiata pero
34	714.5	44.5	975	22	AAH44081	Pinus radiata pero
35	714	44.5	1348	22	AAH44081	Pinus radiata pero
36	710.5	44.3	975	22	AAH44081	Pinus radiata pero
37	708.5	44.2	1177	21	AAH44081	Pinus radiata pero
38	706	44.0	1089	21	AAH44081	Pinus radiata pero
39	699	43.6	1062	24	AAH44081	Pinus radiata pero
40	692.5	43.2	1305	21	AAH44081	Pinus radiata pero
41	692.5	43.2	1305	21	AAH44081	Pinus radiata pero
42	692.5	43.2	1305	21	AAH44081	Pinus radiata pero
43	690	43.0	1065	21	AAH44081	Pinus radiata pero
44	690	43.0	1065	21	AAH44081	Pinus radiata pero
45	687	42.8	1055	10	AAH44081	Pinus radiata pero

ALIGNMENTS

RESULT 1  
AAH44077 standard; CDNA; 1156 BP.

AAH44077; (first entry)

12-SEP-2001

Oryza sativa peroxidase s4235 encoding CDNA seq ID NO:13.

Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;

modification; plant; bacterial infection; Magnaporthe grisea; ss.

Oryza sativa.

Key Location/Qualifiers  
CDS 75..1058  
FT /\*\*tag= a

ET /product= "peroxidase s4235"

XX WO200142475-A1.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000WO-JP08728.

XX 10-DEC-1999; 99JP-0352472.

XX (NORU) JAPAN MIN AGRIC FORESTRY & FISHERIES.

XX Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;

XX Hiraoka S;

XX WPI: 2001-381695/40.

XX P-PSDB: AAB99738.

XX New set of rice peroxidase genes for analysis of peroxidase expression

XX in rice under varying conditions and production of rice plants with

XX desired characteristics

XX Claim 1: Page 152-155; 258bp; Japanese.

XX The present invention describes a set of peroxidase genes found in

XX plants, especially rice, and their homologues, modified forms and

XX fragments, where the sequences of the peroxidase genes in the set are

XX given in AH44071 to AH44091. Also described are: (1) promoters for the

XX control of the gene set; (2) the preparation of cassette vectors using the

XX the genes and promoters; (3) analysis of plant characteristics using the

XX peroxidase set by isolating RNA from the plant, binding the RNA to a

XX membrane, mixing with a labelled peroxidase gene set, incubating, and

XX detecting the label signal to show which genes in the set are expressed

XX in the sample plant; and (4) DNA microarrays for peroxidase gene

XX expression analysis. The set of genes are used for the analysis of the

XX pattern of peroxidase gene expression in particular rice plants and

XX their component tissues and under different environmental conditions,

XX and modification of rice plants to provide desired specificities of

XX peroxidase gene expression to impart particular characteristics to the

XX plants such as response to bacterial infection by Magnaporthe grisea.

XX

XX Sequence 1156 BP; 194 A; 398 C; 364 G; 200 T; 0 other:

XX

Alignment Scores:

Pred. No.: 2,33e-75 Length: 1156

Score: 982.50 Matches: 192

Percent Similarity: 75.00% Conservative: 42

Best Local Similarity: 61.54% Mismatches: 69

Query Match: 61.25% Indels: 9

DB: 22 Gaps: 3

US-10-047-825-4 (1-313) x AH44077 (1-1156)

QY 11 ValAlaValSerLeuLeuSerCysValAlaHisAlaGlnLeuSerProThrPheTyrAla 30

DB 117 GTGGCCCTTGGCGGTGCTCTCGCGCGAGAGCTCACCTGCGCGGGGACTACAC 176

QY 31 SerSerCysProAsnLeuGlnSerIleValArgAlaAlaMetThrGlnAlaValAlaSer 50

DB 177 GCGACGTGCCCCGGGGTGGTGCATCTGCGCGCGGATGCGGAGCACTGCAGAA 236

QY 51 GlnGlnArgMetGlyAlaSerLeuLeuArgLeuPheHisAspCysPheValGlnGly 70

DB 237 GAGTCGCGCATGGCGCGCTCCATCTCCGCTCTTCCAGACAGCTGCTCTCAACGG 296

QY 71 CysAspGlySerIleLeuLeuAsp-----AlaGlyGlyGlnLeuThrAlaGly 86

DB 297 TGGGAGCGCTCCATCTTGTCTGCGAGACGCGCACTTACCGGGGAGAAAGCGCGG 356

QY 87 ProAsnLeuAsnSerValArgGlyPheGlnValIleAspThrIleLeuArgAsnValGln 106

DB 357 CCGAGCGCCAACTCGGTGGCGGGGTACGAGTGTATGACCGCATCAAGCGCAGCTCGAG 416

QY 107 AlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaArgAspGly 126

DB 417 GCCTCCGTCGAGAGCCGACCGCTCTCTCGCGGACATCATCAGCTCGCGCGGACCGCC 476

QY 127 ThrAsnLeuGlnGlyLysProThrTrpSerValProLeuGlyArgArgAspSerThr 146

DB 477 GTCACCTGCTCGGGGCCCGAGACGTGAGCTCCCTGGGGGCGGTACCGCGCGCAG 536

QY 147 AlaSerAlaSerLeuAlaAsnSerAsnProProProProThrAlaSerLeuGlyThrLeu 166

DB 537 ACAGACCGAGAGCGCGCGGAGAACCAACCTGCGCGCGCGCGCGGCGAGCTCGCTCG 596

QY 167 IleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThrAlaLeuSerGlyAla 186

DB 597 CTGTGATGTTTCAAGCGCGCAAGGCGCTGACGCGCGGAGCTCAGCGCTGCGGGCGG 656

QY 187 HisThrIleGlyGlnAlaArgCysThrThrPheArgGlyArgIleTyrGlyAspThrAsp 206

DB 657 CACACCGTCGGGGGCGGCTGCTCCACCTTCCGACGCGCACATCTTACACGACCGCGG 716

QY 207 IleAsnAlaSerPheAlaAla--LeuArgGlnGlnThrCysProArgSerGlyAlaAsp 225

DB 717 GTGACCGCCACCTTCCGCTCGAGCTGCGCGACGACCAAGTCTGCGCGACCGCGCGAG 776

QY 226 GlyAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyrPheThr 245

DB 777 GGCACCTCCGCGCGCGCTGCGCGAGCGCGCCACCACTTCCACCAAGCTTACTTACG 836

QY 246 AsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnGlnLeuPhe----- 261

DB 837 GACCTCTGACCGCGCGCGCTGCTGCTGCGCTGCGCGAGGAGCTTTCGCGAGCGCGCG 896

QY 262 AsnGlySerGlnAspAlaLeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsn 281

DB 897 GGCATGTCACACGAGCGAGCGGTGCTGCGCGCTGCGCGAGCGCGAGCGAGCTTCGCG 956

QY 282 AlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAla 301

DB 957 GCGGACTTGGCGCGGAGTGTAGGCTGGGCAACTGAGCCGCTCACCGGGAGAGAAC 1016

QY 302 GlyGlnIleArgArgAsnCysArgValValAsnSer 313

DB 1017 GCGGAGGTAGCGATCACTGCGCGCGAGTCACTCA 1052

RESULT 2

AAA68187

ID AAA68187 standard; DNA: 1224 BP.

XX AAA68187;

XX 24-OCT-2000 (first entry)

XX Pinus radiata peroxidase nucleotide sequence SEQ ID NO:363.

XX Plant: liginin biosynthetic pathway; Eucalyptus grandis;

XX Pinus radiata; Monterey pine; ds.

XX Pinus radiata.

XX WO200022099-A1.

XX 20-APR-2000.

XX 06-OCT-1999; 99WO-NZ00168.

XX 09-OCT-1998; 98US-0169789.

XX 14-JUL-1999; 99US-0143811.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Bloksberg LN, Havukkala IU;



xx Claim 1; Page 171-174; 258bp; Japanese.

CC The present invention describes a set of peroxidase genes found in  
CC plants, especially rice, and their homologues, modified forms and  
CC fragments, where the sequences of the peroxidase genes in the set are  
CC given in AAH44071 to AAH44091. Also described are: (1) promoters for the  
CC control of the gene set; (2) the preparation of cassette vectors using  
CC the genes and promoters; (3) analysis of plant characteristics using the  
CC peroxidase set by isolating RNA from the plant, binding the RNA to a  
CC membrane, mixing with a labelled peroxidase gene set, incubating, and  
CC detecting the label signal to show which genes in the set are expressed  
CC in the sample plant; and (4) DNA microarrays for peroxidase gene  
CC expression analysis. The set of genes are used for the analysis of the  
CC pattern of peroxidase gene expression in particular rice plants and  
CC their component tissues and under different environmental conditions,  
CC and modification of rice plants to provide desired specificities of  
CC peroxidase gene expression to impart particular characteristics to the  
CC plants such as response to bacterial infection by Magnaporthe grisea.

xx Sequence 1158 BP; 255 A; 322 C; 327 G; 254 T; 0 other;

#### Alignment Scores:

Pred. No.:	6e-73	Length:	1158
Score:	954.50	Matches:	189
Percent Similarity:	71.87%	Conservative:	46
Best Local Similarity:	57.80%	Mismatches:	77
Query Match:	59.51%	Indels:	15
DB:	22	Gaps:	3

US-10-047-825-4 (1-313) x AAH44080 (1-1158)

Oy 1 MetaLaserProThr-----LeuMetGlnCysLeuVal 11  
Db 14 ATGGCTCAGCCAGCTGTCAGCAGAGCGCGCTCAGTGGCCGCTGTGTGTGTG 73  
Oy 12 AlaValSerLeuLeuSerCysValAlaHisAlaGlnLeuSerProThrPheTrAlaSer 31  
Db 74 GTGGTCTCTCCGCGCCGCGCGCGCTGCGCGCAGCTGTCGCCGCTTACTACTGTC 133  
Oy 32 SerCysProAsnLeuGlnSerIleValArgAlaIleMetThrGlnAlaValAlaSerL 51  
Db 134 TCGTGCCTCGGAGGTTCACCGCGGAGCGGGGATGCACCTCGCCATCGCCAGGAG 193  
Oy 52 GlnArgMetGlnAlaSerLeuLeuArgLeuPheHisAspCysPheValGlnGlyCys 71  
Db 194 AAGCGATCGCGCGCTCCATCGCTCTCTTCCAGAGCTGTTGTGTCAGATGTC 253  
Oy 72 AspGlySerIleLeuLeuAsp-----AlaGlyGlyLysThrAlaGlyPro 87  
Db 254 GACGATCTCTCTCTCTGTCGACGACGAGCGAGCTTCACCGCGAGAAAGCGCAACCC 313  
Oy 88 AsnLeuAsnSerValArgGlyPheGluValIleAspThrIleLysArgAsnValGluAla 107  
Db 314 AACAAAGCTCTCGTCAGAGGTTTGAAGTATGACGCGATCAATCAATCGCGGTGAGAC 373  
Oy 108 AlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThr 127  
Db 374 ATCTGCCCGGGGCTGCTCCCTCGCGCAGCATCTCGCATCGCTCCAGGAGACGCTC 433  
Oy 128 AsnLeuGlnGlyProThrTrpSerValProLeuGlyArgArgAspSerThrAla 147  
Db 434 GCCATCTCTGGTGGCGCGCGAGCTGAAGTTGTGCGAGAGATCTCGCGAGCGGCG 493  
Oy 148 SerAlaSerLeuAlaAsnSerAsnProProProProThrAlaSerLeuGlyThrLeuIle 167  
Db 494 AACCTCAGCGGGGCAACAAACATCCCGCGCGAGCTCGGAGCTGCCAACCTTCACC 553  
Oy 168 SerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThrAlaLeuSerGlyAlaHis 187  
Db 554 TCCCTCTTGGCGCGAGCGCTCTCCAGAAAGCATGATGCGCTCTCGGATCTAC 613  
Oy 188 ThrIleGlyGlnAlaArgCysThrThrPheArgGlyArgIleTyrGlyAspThrAspIle 207

Db 614 AACATGGGAGCAGCAGCATCACAAACTTCAGAGCTCATATATACAAACCAACATC 673  
Oy 208 AsnLaserPheAlaAlaLeuArgGlnGlnThrCysProArgSer-----GlyGlyAsp 225  
Db 674 GACAGTGGCTTTGGATGAGAGAGCAATTCAGTTGCTTGTAACACAGATCAGTGAC 733  
Oy 226 GlyAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyrPheThr 245  
Db 734 AATAACTGGACACTCGATCGATCTTCACAGGCCCAACGCTGTGAGAAACAACTACTCAAG 793  
Oy 246 AsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnLeuLeuPheAsnGlyGlySer 265  
Db 794 AACCTCTCTGTCAGAAAGAGGCGCTCCGATTCGATCAGAGCTCTCAATGCGGAGCC 853  
Oy 266 GlnAspLeuLeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsnAlaAspPheVal 285  
Db 854 ACTATGCTCTTGTTCAGTCTTACATTAAGTACAGCAGCATCTTTGGGATTTGTG 913  
Oy 286 AlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArg 305  
Db 914 ACGGCAATGATCAAGATGGGAGCATCACCGTTGACGGATCAACAGGAGATCAG 973  
Oy 306 ArgAsnCysArgValValAsn 312  
Db 974 AAGAACTGCAGAAAGATTAAT 994  
RESULT 4  
AAA68193  
ID AAA68193 standard; DNA; 1171 BP.  
AC  
XX AAA68193:  
XX  
XX 24-OCT-2000 (first entry)  
DT  
XX  
DE pinus radiata peroxidase nucleotide sequence SEQ ID NO:369.  
XX  
KW Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;  
KM pinus radiata; Monterey pine; ds.  
XX  
OS pinus radiata.  
XX  
XX WO200022099-A1.  
PN  
XX 20-APR-2000.  
PD  
XX  
XX 06-OCT-1999; 99MO-N200168.  
PF  
XX  
XX 09-OCT-1998; 98US-0169789.  
PR  
XX 14-JUL-1999; 99US-0143811.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
XX Bloksberg LN, Havukkala IJ;  
XX  
XX WPI; 2000-317962/27.  
XX  
XX Novel polynucleotide encoding enzymes involved in lignin-biosynthetic  
PT pathway useful for producing transgenic plants especially eucalyptus  
PT and pine species having altered lignin content, composition and  
PT structure  
XX  
PS Claim 1; Page 189-190; 213pp; English.  
XX  
XX The present invention describes isolated polynucleotides and proteins  
XX encoding and representing the enzymes cinamamate 4-hydroxylase (C4H),  
XX coumarate 3-hydroxylase (C3H), phenolase (PND), O-methyl transferase  
XX (OMT), cinammyl alcohol dehydrogenase (CAD), cinamoyl-CoA reductase  
XX (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate-CoA ligase (CL),  
XX coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),  
XX laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,  
XX caffeic acid methyl transferase, caffeoyl CoA methyl transferase,

CC coumarate CoA ligase, cytochrome P450 1X1A, diphenol oxidase, flavanol  
 CC glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,  
 CC which are involved in the lignin biosynthetic pathway. The  
 CC polynucleotides can be used for modulating lignin content, lignin  
 CC composition and the structure of a plant, especially eucalyptus and pine  
 CC species, and for modifying the activity of an enzyme involved in lignin  
 CC biosynthetic pathway, and for producing a plant having altered lignin  
 CC content, composition and structure. They can be used for designing probes  
 CC and primers useful for detecting similar DNA and RNA sequences in any  
 CC organism and for PCR amplification. The lignin content can be efficiently  
 CC modified using the polynucleotides. AAB67908 to AAB68201 and AAB16341 to  
 CC AAB16449 represent polynucleotide and protein sequences used in the  
 CC exemplification of the present invention.

XX Sequence 1171 BP: 329 A; 282 C; 263 G; 297 T; 0 other;

#### Alignment Scores:

Pred. No.:	1,48e-72	Length:	1171
Score:	950.00	Matches:	186
Percent Similarity:	73.25%	Conservative:	44
Best Local Similarity:	59.24%	Mismatches:	74
Query Match:	59.23%	Indels:	10
DB:	21	Gaps:	3

US-10-047-825-4 (1-313) x AAA68193 (1-1171)

OY 9 CysLeuValAlaValSerLeuLeuSerCys-----ValAlaHisAlaGlnLeu 24  
 DB 70 TGTATAGCTGTATAGCTGTATATCTGTTCTATTACAGATCTGTAATGGGCGAGCTG 129  
 OY 25 SerProThrPheYrAlaSerSerCysProAsnLeuGlnSerIleValAlaAlaMet 44  
 DB 130 AGCTGACAGTTTTCAGCAAAATCTTCCCGACGGACCTGCGTATGTAAGCGCGAGTG 189  
 OY 45 ThrGlnAlaValAlaSerGlnGlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHis 64  
 DB 190 AAGCAGAGCGCTGCTACAGAAACGATGGTGGCTCTTGGCCGCTCAGCTTCTAC 249  
 OY 65 AspCysPheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGly----- 80  
 DB 250 GACTGCTCTGTTATGTTGCGATGGTCCGTTCTGTTGGACGATCTTCAGCATTACT 309  
 OY 81 GlyGluYrThrAlaGlyProAsnLeuAsnSerValAlaGlyPheGlnValIleAspThr 100  
 DB 310 GCGGAGAGACAGCTAATCCCAATCTCTGCGAGGAGCTTCAGAGTATATGATACC 369  
 OY 101 IleYrAlaGlnValAlaGlnAlaAlaCysProGlyValAlaSerCysAlaAspIleLeuAla 120  
 DB 370 ATAAAGAGCAATGTGAGAAACCTTGACAGTGCCTTCTGCGACAGATCTCTGCC 429  
 OY 121 LeuAlaAlaArgAspGlyYrThrAsnLeuLeuGlyGlyProThrTrpSerValProLeuGly 140  
 DB 430 ATTGCTGCTCGTATCTGTTGTTGAACCTGGCGGCTCTTCAATGACAGTATGTTGGCA 489  
 OY 141 ArgArgAspSerThrThrAlaSerIleLeuAlaAsnSerAsnProProProThr 160  
 DB 490 AGCGGAGACTCGAACAGCTAGCAAAAGCGGTGCAAAACAGTAAATATTCGCGCTCGACT 549  
 OY 161 AlaSerLeuGlyYrThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMet 180  
 DB 550 TCCAGTCTGAGCAACCTCTCTCACTATTCACAGCGAGGCTCTCGCAAGGAAGATG 609  
 OY 181 ThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArgIlyArg 200  
 DB 610 GTTGCACTTTCGCGGCTATACATCGGCGAGGCGCATGCAAAATTCAGACCCAT 669  
 OY 201 IleYrGlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysPro 220  
 DB 670 ATTACACAGACAGCAACATAGACAGTGCCTGACGCCACTTCCTGCAAGAGTGC 729  
 OY 221 ArgSer-----GlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrProValArg 238  
 DB 730 AGTACCAAGAGCTCGGAGACAGCACTTGTGCGCATTTGATATATGACTCCACAGTGTG 789

OY 239 PheAspThrAlaYrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAspGln 258  
 DB 790 TTTCAGCAAAATCTATACAGCGACCTCGAAAGCCAAAAGGACTTCTCCACCGACAG 849  
 OY 259 GluLeuPheAsnGlyIleSerGlnAspAlaLeuValArgIleYrSerAlaSerAlaSer 278  
 DB 850 GAACCTCTCAAGGAGGAGCTCCACTGATTCACAGGTACTAGTACGCCCTCCAAACGAGAG 909  
 OY 279 LeuPheAsnAlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThr 298  
 DB 910 ACCCTTCTCCGATTTTCCTGCGGCAATGTTAAAGATGGAATATCAAACTCTTACC 969  
 OY 299 GlyThrAlaGlyGlnIleArgArgAsnCysArgValValAsn 312  
 DB 970 GGCACCAAGCGGACAGATCCCAAGAACTCGAGAACCCAAAC 1011

RESULT 5  
 AAT95782  
 ID AAT95782 standard; cDNA to mRNA; 1144 BP.  
 XX  
 AC AAT95782:  
 XX  
 DT 27-APR-1998 (first entry)  
 XX  
 DE Stylosanthes humilis Shpx6 peroxidase cDNA.  
 XX  
 KW Peroxidase; Shpx6; transgenic plant; fungus resistance;  
 KW disease protection; Phytophthora parasitica;  
 KW Leptosphaeria maculans; Sclerotinia sclerotiorum; ds.  
 OS Stylosanthes humilis strain Paterson.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 42..1004  
 FT sig\_peptide /tag- a  
 FT /tag- b  
 FT mat\_peptide 114..1001  
 FT /tag- c  
 FT  
 PN W09741237-A1.  
 XX  
 PD 06-NOV-1997.  
 XX  
 PF 29-APR-1997; 97MO-AU00253.  
 XX  
 PR 29-APR-1996; 96AU-0009532.  
 XX  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI (GRAI-) GRAINS RES & DEV CORP.  
 XX  
 PI Goulter KC, Kazan K, Manners JM;  
 XX  
 DR WPI, 1997-549739/50.  
 DR P-PSDB; AAW38217.  
 XX  
 PT Plant cell transformed with Stylosanthes humilis Shpx6 peroxidase  
 PT useful to confer on plants resistance to fungi, e.g. Phytophthora  
 PT parasitica, Leptosphaeria maculans or Sclerotinia sclerotiorum  
 XX  
 PS Claim 3; Page 14-15; 30pp; English.

This cDNA sequence includes a coding region for the 320-amino acid  
 Shpx6 peroxidase (see AAW38217) of Stylosanthes humilis. A claimed  
 method of engineering a plant to fungal resistance comprises  
 introducing into cells of the plant a DNA construct comprising a  
 promoter that is constitutively operative in the plant cell  
 (preferably the cauliflower mosaic virus 35S promoter) and the  
 Shpx6 sequence, or a hybridizing sequence or fragment that encodes  
 an enzyme with peroxidase activity. Also claimed are plant cells  
 harboring the DNA construct, a plant comprising such cells, and  
 material from such a plant, especially seed, pollen, a stem segment





Score: 926.00 Matches: 185  
 Percent Similarity: 72.52% Conservative: 42  
 Best Local Similarity: 59.11% Mismatches: 82  
 Query Match: 57.73% Indels: 4  
 DB: 22 Gaps: 3

US-10-047-825-4 (1-313) x AAH4085 (1-1310)

QY 2 AlasSerProthrleuMetGlnCysLeuValAlaValSerleuLeuSerCysValAlaHis 21  
 DB 90 GCTTCTTCTGACGCTTAATAGCTGCTGCTGCAAGCTATAGGCTCA--GCCGCTCG 146  
 QY 22 AlAGlnLeuSerProthrPheThrAlaSerSerCysProkmsLeuInSerIleValArg 41  
 DB 147 GCCGAGCTGCGGCGAGCTTCTACGACACCTGCTGCCCAACCATTTGTCACCATCAG 206  
 QY 42 AlAlaMetThrGlnAlaValAlaSerGlnArgMetGlnAlaSerleuLeuArgLeu 61  
 DB 207 AGCGAGTACAGCGCGCGTGAACAGGAGCCCGAATGAGCACTCGCTGTCAGAGCTG 266  
 QY 62 PheMetHisPcySerPheValGlnGlnGlnCysAspGlySerIleLeuLeuAspAlaGly 81  
 DB 267 CACTTCACACAGCTGCTGTCGCAAGGCTGTGACGCTGCTGCTG--TCCGCGCAG 323  
 QY 82 GluIysThrAlaGlyProAlaLeuAsnSerValArgGlyPheGlnValIleAspThrIle 101  
 DB 324 GAGCAGAAATGACAGCGCGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383  
 QY 102 LysArgAsnValGlnAlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeu 121  
 DB 384 AAGACGAGGTGCGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 443  
 QY 122 AlaAlaArgAspGlyThrAlaLeuLeuGlnGlyIleProThrTrpSerValProLeuGlyArg 141  
 DB 444 GCGCGCGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 503  
 QY 142 ArgAspSerThrThrAlaSerAlaSerleuAlaAsnSerAnProProProThrAla 161  
 DB 504 AGGAGCTCCACACATGCGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 563  
 QY 162 SerleuGlnThrleuIleSerleuPheGlnArgGlnGlnIleuSerProAlaGaspMetThr 181  
 DB 564 AGCCTCCCTGAACCTATGCGCAATTTCTCCAGAAAGGAGCTGACAGTACCGCATGTT 623  
 QY 182 AlaleuSerGlnAlaHisThrIleGlnAlaAlaArgCysThrThrPheArgGlyArgIle 201  
 DB 624 GCTTCTCAGGCGCAGACAGATCGGCGAGCGCATGTCAGAACTTCAAGGAGAGGCTC 683  
 QY 202 TyrGlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArg 221  
 DB 684 TACAACGACAAACATCGACTCCAGCTCGCGAGCGGCTCAAGGCCAACCTGCCACG 743  
 QY 222 -----SerGlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrProValArgPhe 239  
 DB 744 CCGAGCGGACCGCGGCGAGCAACCTGCGCGCTGAGCAGCAGCAGCGCCCAACGCGCTTC 803  
 QY 240 AspThrAlaThrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnGlu 259  
 DB 804 GACGCGCGCTTACTACCAACCTGCTCTCCACACAGGCGCTCTGCGACTCGGACGAGT 863  
 QY 260 LeuPheAsnGlyGlySerGlnAspAlaLeuValArgGlnThrSerAlaSerAlaSerLeu 279  
 DB 864 CTCTTCAACGCGCGGCGAGCAACAGCGTGAAGAACTTCTGCTCAACAGCGCGCG 923  
 QY 280 PheAsnAlaAspPheValAlaAlaMetIleArgMetGlnValGlyValIleuThrGly 299  
 DB 924 TTCAACGCGCGCTTCAAGCGCGCTGTCAGATGGGCAACATCTCGCCCTTGACTGGA 983  
 QY 300 ThrAlaGlnIleArgArgAsnCysArgValValAsn 312  
 DB 984 ACCCAGGCGCGATCAGGCTCACTGCTCCAGGTTAAC 1022  
 RESULT 7

AAH4082  
 ID AAH4082 standard; cDNA; 1171 BP.  
 XX  
 AC AAH4082;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Oryza sativa peroxidase s10927 encoding cDNA SEQ ID NO:23.  
 XX  
 KM Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;  
 XX modification; plant; bacterial infection; Magnaporthe grisea; ss.  
 XX  
 OS Oryza sativa.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 53..1033  
 FT /tag= a  
 FT /product= "peroxidase s10927"  
 PN  
 PN W0200142475-A1.  
 PD 14-JUN-2001.  
 PF 08-DEC-2000; 2000WO-JP08728.  
 PR 10-DEC-1999; 99JP-0352472.  
 XX  
 PA (NORU) JAPAN MIN AGRIC FORESTRY & FISHERIES.  
 XX  
 PI Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;  
 PI Hiraga S;  
 PI  
 DR WPI: 2001-381695/40.  
 DR P-PSDB: AAB99743.  
 XX  
 PT New set of rice peroxidase genes for analysis of peroxidase expression  
 PT in rice under varying conditions and production of rice plants with  
 PT desired characteristics -  
 PS Claim 1; Page 183-186; 258pp; Japanese.  
 CC The present invention describes a set of peroxidase genes found in  
 CC plants, especially rice, and their homologues, modified forms and  
 CC fragments, where the sequences of the peroxidase genes in the set are  
 CC given in AAH4071 to AAH4091. Also described are: (1) promoters for the  
 CC control of the gene set; (2) the preparation of cassette vectors using  
 CC the genes and promoters; (3) analysis of plant characteristics using the  
 CC peroxidase set by isolating RNA from the plant, binding the RNA to a  
 CC membrane, mixing with a labelled peroxidase gene set, incubating, and  
 CC detecting the label signal to show which genes in the set are expressed  
 CC in the sample plant; and (4) DNA microarrays for peroxidase gene  
 CC expression analysis. The set of genes are used for the analysis of the  
 CC pattern of peroxidase gene expression in particular rice plants and  
 CC their component tissues and under different environmental conditions,  
 CC and modification of rice plants to provide desired specificities of  
 CC peroxidase gene expression to impart particular characteristics to the  
 CC plants such as response to bacterial infection by Magnaporthe grisea.  
 XX  
 SQ Sequence 1171 BP; 172 A; 429 C; 362 G; 208 T; 0 other.  
 Alignment Scores:  
 Pred. No.: 2,446-68 Length: 1171  
 Score: 901.00 Matches: 170  
 Percent Similarity: 73.57% Conservative: 61  
 Best Local Similarity: 54.14% Mismatches: 77  
 Query Match: 56.17% Indels: 6  
 DB: 22 Gaps: 3  
 US-10-047-825-4 (1-313) x AAH4082 (1-1171)  
 QY 6 LeuMetGlnCysLeuValAlaValSerLeu---LeuSerCysValAlaHisAlaGlnLeu 24  
 DB 89 CTGCTCTGCTCTCTCTCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 148

Oy	25	serProthrethrylaIaIaSerSerCysProAsnLeuInserIleValArgAlaIaMet	44		
Db	149	TCGCGCGGGGTTCACTCGGCGAGCTGCCACCGGACGGCGCTCGCGGCGAGTCAAG	20		
Oy	45	ThriGlnAlaValAlaSerGlnInArgMetGlyAlaSerLeuLeuArgLeuPheHis	64		
Db	209	TCGCAGGCGGTATGAAACACACGAGCGCGCGGCGCGCGCTCTTATACAC	26		
Oy	65	AspCysPheValGlnGlyAspGlySerIleLeuLeuAsp-----AlaGly	80		
Db	269	GACTGCTTCGTCGCGCGCGCACCGCTCGGAGTCTCTGCACGACACCCCGCGGCGCC	32		
Oy	81	GlyGlnIuSerThrIaGlyProAsn---LeuAsnSerValArgGlyPheGlnValAsp	99		
Db	329	GCGGAGAAAGGCGCTCGCGCCCAACCGCTCGCTGCAGACCGCTTACGCTCGTCAC	38		
Oy	100	ThrIleLysArgAsnValGlnAlaIaCysProGlyValValSerCysAlaAspIleLeu	11		
Db	389	ACCATCAAGGCCCGAGGCTGAGGCGCTGCGGCCCCACCGTCTCTGCGCGAGCTCTC	44		
Oy	120	AlaLeuAlaIaIaArgspolYThrAsnLeuLeuGlyIleProThrIrpSerValProLeu	13		
Db	449	GCAATCGCGCGCGCGCACAGCTCAACCTGTCGCGGCGGAGCTGGGCGGCTCGCTC	50		
Oy	140	GlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProPro	15		
Db	509	GCGCGCGCGAGCGCGCTGCGCGCGAGTGGAGCCGCGGTCTGCACGACTCCCGGCGCC	56		
Oy	160	ThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAsp	17		
Db	569	GAGCGCGACATCCCGCGCTCTCCGCTTCCGCTTCCCGCCCAAGGAGGCTTGAGCTCGCGAC	62		
Oy	180	MetThrAlaLeuSerGlyAlaHisThrIleGlyIlnAlaArgCysThrThrPheArgGly	19		
Db	629	CTGCGCGCGCTGTCGCGCGCGCACCGCTGCGCGCGCGCGCGCGCTCAACTTCGCGAC	68		
Oy	200	ArgIleArgGlyAspThrAsnIleAsnAlaSerPheAlaIaLeuArgGlnGlnThrCys	21		
Db	689	CGGCTACAGCGAGCGCCACAGCTGAGCCCGCGCTTCGCTCGCACGACGAGCGAGTCTTC	74		
Oy	220	ProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrProValArgPhe	23		
Db	749	CCGCGCGCTCGCGCGCGACCCCGCTGCGCGCTGAGCTCCCTGACCCCGCACCGCTTC	80		
Oy	240	AspThrAlaIleArgPheThrAsnLeuSerArgArgGlyLeuPheHisSerAspGlnGly	25		
Db	809	GACACAGCGCTACTACCGCACTCTGTCGCCCGCGCGCGGCTGTCGACTTCGACAGAG	86		
Oy	260	LeuPheAsnGlyGlySerGlnAspAlaLeuValArgGlnIleSerAlaSerLeu	27		
Db	869	CTGTTCAACACAGCGCGCGGTGACTCGGTGTGTCAGCTGACAGCTCAACGCCCGCGC	92		
Oy	280	PheAsnAlaAspPheValAlaIaMetIleArgMetGlyAsnValGlyValLeuThrGly	29		
Db	929	TTCCTCTCGGACTTTCGCGCGCTCAATGTCAGGCTCGGGAACATCGCGCTTGACCGGG	98		
Oy	300	ThrAlaGlyGlnIleArgArgAsnCysArgValValAsnSer	313		
Db	989	TCAACCGCGGAGGTGAGGCTCACTGACGAGGAAGTAATTC	1030		
RESULT 8					
ID	AAA68176	AAA68176 standard; DNA; 1391 BP.			
XX	AAA68176;				
AC	XX				
XX	XX				
DT	XX				
24-OCT-2000	(first entry)				
XX	XX				
DE	Eucalyptus grandis peroxidase nucleotide sequence SEQ ID NO:352.				
XX					
Plant:	Lignin; Lignin biosynthetic pathway; Eucalyptus grandis;				
XX	Pinus radiata; Monterey pine; ds.				

	XX	Eucalyptus grandis.
	O5	
	XX	WO200022099-A1.
	PN	
	XX	
	PD	20-APR-2000.
	XX	
	PF	06-OCT-1999; 99WC-NZ00168.
	XX	
	--PR	09-OCT-1998; 98US-0169789.
	PR	14-JUL-1999; 99US-0143811.
	PA	(GENE-) GENESIS RES & DEV CORP LTD.
	PA	(FLET-) FLETCHER CHALLENGE FORESTS LTD.
	PI	Bloksberg LN, Havukala IJ;
	DR	
	XX	WPI: 2000-317962/27.
	PT	
	PT	Novel polynucleotide encoding enzymes involved in lignin-biosynthetic pathway useful for producing transgenic plants especially eucalyptus and pine species having altered Lignin content, composition and structure -
	XX	
	F5	Claim 1; Page 184; 213pp; English.
	CC	The present invention describes isolated polynucleotides and proteins encoding and representing the enzymes cinnamate 4-hydroxylase (C4H), coumarate 3-hydroxylase (C3H), phenolase (PMU), o-methyl transferase (OMT), cinamoyl alcohol dehydrogenase (CAD), cinamoyl-CoA reductase (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL), coniferol glucosyl transferase (CGT), coniferyl beta-glucosidase (CBG), laccase, peroxidase, ferulate-5-hydroxylase (FSH), alpha-amylase, caffeic acid methyl transferase, caftaroyl CoA methyl transferase, coumarate CoA ligase, cycochrome P450 IXIA, diaphenol oxidase, flavanol glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase, which are involved in the lignin biosynthetic pathway. The polynucleotides can be used for modulating lignin content, lignin composition and the structure of a plant, especially Eucalyptus and Pine species, and for modifying the activity of an enzyme involved in lignin biosynthetic pathway, and for producing a plant having altered lignin content, composition and structure. They can be used for designing probes and primers useful for detecting similar DNA and RNA sequences in any organism and for PCR amplification. The lignin content can be efficiently modified using the polynucleotides. AA61908 to AAA68201 and AAB16341 to AAB16449 represent polynucleotide and protein sequences used in the exemplification of the present invention.
SQ	Sequence 1391 BP; 399 A; 321 C; 307 G; 364 T; 0 other:	
	Alignment Scores:	
	Pred. NO.:           4.09e-68         Length:         1391	
	Score:              899.50           Matches:          180	
	Percent Similarity: 71.52%          Conservative:     46	
	Best Local Similarity: 56.96%       Mismatches:       85	
	Query Match:        56.08%	Indels:           5
		Gaps:            2
	Df:                  21	
US-10-047-825-4 (1-313) x AAA68176 (1-1391)		
OY	1 MetAlaserProthrleuMetgIncyLeuValAlaIalSerleuLenuSerCySyaValAla	20
	:      :	
Db	246 ATGTGGCGCTTTTCGTCGTCGTGCCTTGCCACACTGCCATTATACC---ACTGCC	302
OY	21 HisAlaclnleuSerProthrPhetrAlaserSerCySProAsnleuGlInsetlleVal	40
	:::        ::::	
Db	303 CGTGTAGCTCTCACGAGETCATTCATCAACAATGCCAAAGCATTTGCGATTGTT	362
OY	41 ArgAlaIlaMetThrnglnIaValAlaIsaruglnInargMeGIyAlaSerleuLenArg	60
	:	
Db	363 GCAGCTGACGATACAAAAGAATCAAGAAATGAAGACC GGAGGGGCGCTCTTGGCTTGG	422
OY	61 Leuphehenlaspcysphevalinglnglcysaspqglyserillelueneulaaspalagly	80

```

DB 423 CTCACCTTCATGCTCTGTCATGAGTGGTGGATGCGCTGCATATTGTTGGATGACAGC 482
QY 81 -----GlycyluylsThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlu 96
DB 483 CCTAGCTCTGCTGGCGAAGAACAGAGCTCCGAAACAAATTCCTCGAGAGGTTCCGA 542
QY 97 ValIleAspThrIleLysArgAsnValGluAlaIleCysProGlyValAlaSerCysAla 116
DB 543 GTGATGACCGCATCAAGGTAGTCTGGAGAGAGAGTGGCTTGGAGTGGTCTCTGTGCA 602
QY 117 AspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGlyProThrTrpSer 136
DB 603 GATATGCTGCTCCCTGCTGCTCCGACATCAGCTTCATTTGGAGAGTCTTCATTCAGACC 662
QY 137 ValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnPro 156
DB 663 GTTAGCTTAGGAGAGAGATTCATTAAGTCTAGAGAGAGCTTGTCTAACCTCCATATA 722
QY 157 ProProProThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSer 176
DB 723 CCTGACCTCTCTTAATCTATGCTGCTCATACCAAGCTTGGCTGCTCAGGTTCTTCA 782
QY 177 ProArgAspMetThrAlaLeuSerGlyAlaIleHisThrIleGlyGlnAlaArgCysThrThr 196
DB 783 GTCAGAAACATGTGGCTCTTCTGCTTCATCATTACCATTCGCTTACGAGATGCACTTCC 842
QY 197 PheArgGlyArgIleThrGlyAspThrAspIleAsnAlaSerPheAlaLeuArgGln 216
DB 843 TTCCGAGACGAGATCTACACAGCTGCAACATAGATACATCTTCCGCTCAATAATTGCG 902
QY 217 GlnThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrPro 236
DB 903 AAGATATGTCCTGAGATTTGGAATGATGATGCTCTCAAGGCTACATCCAAAGCGG 962
QY 237 ValArgPheAspThrAlaIleThrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSer 256
DB 963 ACCTTCTTTGACAACTTCTACTACCAATTTACTGCAGAGAGAGGCTTCTTCACTCT 1022
QY 257 AspGlnGlnLeuPheAsnGlnGlySerGlnAspAlaLeuValArgGlnThrSerAlaSer 276
DB 1023 GATCAAGAGCTCTTCATAGGCACTCTGTGATTCAGTGTCAAGAGATGATGATCGGAC 1082
QY 277 AlaSerLeuPheAsnAlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyVal 296
DB 1083 ACAGAAATTTTTCGAGATTTTGGCAAGGCAATGATCAATAATGCGAAATTAAGCCC 1142
QY 297 LeuThrGlyThrAlaGlyGlnIleArgArgAsnGlyAsnValValAsn 312
DB 1143 CCCAAAGAGAGCAATGCTCAATTAAGAAATTCGACGAAAGTGAAC 1190
RESULT 9
AAF90234
ID AAF90234 standard; cDNA; 1379 BP.
XX
XX AAF90234;
XX
XX 06-AUG-2001 (first entry)
XX
XX Nucleotide sequence of a peroxidase P7X.
XX
XX Peroxidase; P7X gene; pathogen defence; nematode resistance; ss.
XX
XX Zea mays.
XX
XX WO200138485-A2.
XX
XX 31-MAY-2001.
XX
XX 24-NOV-2000; 2000MO-US30159.
XX
XX 24-NOV-1999; 99US-0167229.
XX

```

```

PA (UMIS ) UNIV MISSISSIPPI STATE.
XX
XX Padelimas LS, Reichert NA:
XX
XX WPI: 2001-355920/37.
XX
XX Novel gene encoding peroxidase P7X protein, and its promoter, useful
XX for producing transgenic plants that are resistant against nematode
XX infections
XX
XX Claim 1; Fig 3; 34pp; English.
XX
XX The present sequence represents the cDNA sequence of a peroxidase
XX P7X gene. The P7X gene is isolated from maize inbred line Mp307.
XX Peroxidases reduce hydrogen peroxide or molecular oxygen in the
XX presence of an electron donor. Plant peroxidases are involved in
XX pathogen defence responses. DNA constructs or transcription cassettes
XX comprising peroxidase P7X gene and its promoter are useful for
XX providing nematode resistance in plants.
SQ Sequence 1379 BP; 322 A; 420 C; 370 G; 267 T; 0 other;
Alignment Scores:
Pred. No.: 7.33e-68 Length: 1379
Score: 896.50 Matches: 182
Percent Similarity: 71.83% Conservative: 50
Best Local Similarity: 56.35% Mismatches: 79
Query Match: 55.89% Indels: 12
DB: 22 Gaps: 5
US-10-047-825-4 (1-313) x AAF90234 (1-1379)
QY 1 MetAlaSerProThrLeuMetGlnCysLeuValAlaValSerLeuSerCysVal---19
DB 72 ATGGCGGCTCTGTTTCTGCTCTTGCCTTATTTACCTGCTGCTGGCGGTGCTG 131
QY 20 -----AlaHisAlaGlnLeuSerProThrPheThrAlaSerCys 33
DB 132 GTGGCGCTGGCGGCGCGCGTGGCGAGCTGTGCTGAGCTTCTACAGACGTCATGC 191
QY 34 ProAsnLeuGlnSerIleValAlaArgAlaAlaMetThrGlnAlaValAlaSerGlnArg 53
DB 192 CCCAAGGCGCTGTCACATCAGAGCGCGCTGAACCTCCGGGTGAGGACAGACCTTCGC 251
QY 54 MetGlyAlaSerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCysAspGly 73
DB 252 GTGGGGGCTGCTGCTGCTGAGGCTCATTTCCAGCACTGCTTTGCTGGGGCTGGACGCG 311
QY 74 SerIleLeuLeu---AspAlaGlyGlyGlyuylsThrAlaGlyProAsnLeuAsn---Ser 91
DB 312 TCCCTTGTGCTGAACGACGACGTCAGAGGAGCAGCGGCCGGAATCTAATCTGAC 371
QY 92 ValArgGlyPheGluValIleAspThrIleLysArgAsnValGluAlaIleCysProGly 111
DB 372 CCAAGGGCTTCTGCTGTCAGACGATCAAGAGCGCAGTGTGAGTGTGCTGCGGCGG 431
QY 112 ValIleSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGly 131
DB 432 ATGCTCTCTGCGCGCGACATCTCGCTGGCGCGCAGGAGCGAGATGCTGCTGGC 491
QY 132 GlyProThrTrpSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeu 151
DB 492 GGGCCTTGTGAGACGTTCTGTAGGCGGAGAGGAGCTCT---ACCGTCTATTCACAGA 548
QY 152 AlaAsnSerAsnProProProProThrAlaSerLeuGlyThrLeuIleSerLeuPheGly 171
DB 549 CAGACACGACGACCCCGACCTCCGACGCTAGCCCTCCACAGACTTTTGTCTGCTAATAC 608
QY 172 ArgGlnGlyLeuSerProArgAspMetThrAlaLeuSerGlyAlaIleHisThrIleGlyGln 191
DB 609 AAGAAATCTCAACCCACAGCATGTGCTGCTCTCTAGAGAGCTCACACATGCGACAG 668
QY 192 AlaArgCysThrThrPheArgGlyArgIleThrGlyAspThrAspIleAsnAlaSerPhe 211

```

Db	669	GGCGAGTCTGCGACCTTCAACGACCAATCTACACAGCACACCAACATCAACTCGGCTTC	728
Oy	212	AlAlAlAlAlAeuArvgInglInThrCysProArGserGlyGlyAspGlyAsnLeuAlaProIle	231
Db	729	GCGGGCTGCTGCACAGGGCCCAACTGCCCCAGGGAGGACACCGCC--CTTGGCGCGCTG	785
Oy	232	AspValGlnThrProValArgPheAspThrAlaIArgPheThrAsnLeuLeuSerArgArg	251
Db	786	CACACACACGACGCCCCAACCGCTTGACACAGCCCTACTACACCAACTGCTGTCCAGAA	845
Oy	252	GlyLeupheHisSerAspGlnGluLeupheAsnGlyGlySerGlnAspAlaLeuValArg	271
Db	846	GGCGCTCGCAGCTCGGACACGACGACTCTTCAACAAAGGAGACACCGCACACGCTCAG	905
Oy	272	GlnTrSerAla-SerAlaSerLeupheAsnAlaAspPheValAlaIAlaMetIleArg	291
Db	906	AGCTTCGGGCTCCACGACGCTGCGCTTCAACAGGCGCTTGCCACGCGCAATGGTCAAG	965
Oy	291	tGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsnCysArgVal	311
Db	966	GGGCACTCAGCCGCCAGACGCGAACCAGGCGGCAATCATCGGCGCACTGCTGAAG	1025
Oy	311	IAsnSer 313	
Db	1026	CACCTCG 1032	
RESULT 10			
AAH44083			
ID	AAH44083	standard; cDNA: 1223 bp.	
XX	AAH44083;		
XX	12-SEP-2001	(first entry)	
DE	Oryza sativa	peroxidase s14493 encoding cDNA SEQ ID NO:25.	
XX	Oryza sativa	rice; peroxidase; POX; characteristic; gene expression; modification; plant; bacterial infection; Magnaporthe grisea; ss.	
XX	Oryza sativa.		
XX	Key	Location/Qualifiers	
FT	CDS	20..982	
FT		/tag- a	
FT		/product- "peroxidase s14493"	
XX	WO200142475-A1.		
XX	14-JUN-2001.		
XX	08-DEC-2000; 2000MO-JP08728.		
XX	10-DEC-1999; 99JP-0352472.		
XX	(NORU ) JAPAN MIN AGRIC FORESTRY & FISHERIES.		
XX	Ohashi Y, Mitsuhashi I, Sasaki T, Nagamura Y, Ito H, Iwai T;		
XX	Hiraga S;		
XX	WPI: 2001-381695/40.		
XX	P-PSDB: AAB99744.		
XX	New set of rice peroxidase genes for analysis of peroxidase expression		
XX	in rice under varying conditions and production of rice plants with		
XX	desired characteristics		
XX	Claim 1; Page 189-192; 258bp; Japanese.		

CC control of the gene set; (2) the preparation of cassette vectors using  
CC the genes and promoters; (3) analysis of plant characteristics using the  
CC peroxidase set by isolating RNA from the plant, binding the RNA to a  
CC membrane, mixing with a labelled peroxidase gene set, incubating, and  
CC detecting the label signal to show which genes in the set are expressed  
CC in the sample plant; and (4) DNA microarrays for peroxidase gene  
CC expression analysis. The set of genes are used for the analysis of the  
CC pattern of peroxidase gene expression in particular rice plants and  
CC their component tissues and under different environmental conditions,  
CC and modification of rice plants to provide desired specificities of  
CC peroxidase gene expression to impart particular characteristics to the  
CC plants such as response to bacterial infection by *Magnaporthe oryzae*.  
XX Sequence 1223 BP; 234 A; 397 C; 338 G; 254 T; 0 other;

**Alignment Scores:**

Pred. No.:	5,576-67	length:	1223
Score:	885.50	Matches:	179
Percent Similarity:	72.1%	Conservative:	46
Best Local Similarity:	57.37%	Mismatches:	76
Query Match:	55.21%	Indels:	11
DB:	22	Gaps:	4

US-10-047-825-4 (1-313) X AAH44083 (1-1223)

QY	11	ValAlaValSerLeu-----LeuSerCysValAlaHisAlaGlnLeuSerPro	26
Db	47	GGGGGGGAGACGTTTGGTGGTGCGCTTGGGGCGGGCGGCGTCCAGCTGTCCAGC	106
QY	27	ThrPheTyrAlaSerSerCysProAsnLeuGlnSerIleValAlaArgAlaMetThrGln	46
Db	107	ACGTTCTACGGCTGTGTGTGGCCCGACCGCGGTGTCAGACGATTCAGAGACGGCGTGAACCGC	166
QY	47	AlaValAlaSerGlnGlnArgMetCylAlaSerLeuLeuArgLeuPhePheHisAspCys	66
Db	167	GGGGGGCCAGGAGACCCCGGATGGGGCCCTCCCTGCTCAGAGCTCCACTTCCAGACTGC	226
QY	67	PheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGly-----GlyGln	82
Db	227	TTTGTCCAAAGATGGGAGCGGTGCTGATCTGCTGGCCGACAAATGCCACTTCCGGGGGAG	286
QY	83	LysThrAlaGlyProAsnLeuAsnSerValArgGlyPheGluValIleAspThrIleLys	102
Db	287	CAGGGGTGGTTCCTTAATGTCATACTCGCTGAGGGGATTTCAGAGCTCATCTCTACGATTAAAG	346
QY	103	ArgAsnValGlnAlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAla	122
Db	347	ATGCAACTCGAGGACATCTTTCGAGGAGACGCTTCCTGAGCGGCACATCTTGGTGTGCC	406
QY	123	AlaArgAspGlyThrAsnLeuLeuGlyGlyProThrTrpSerValProLeuGlyArgArg	142
Db	407	GCCCGCCGACTCGTCGTGCGCCCTAGGAGTCTCAATCCCGGTGAGCTCGGAGAGAGG	466
QY	143	AspSerThrThrAlaSerAlaSerIleuAlaSerAsnSerAsnProProProThrAlaSer	162
Db	467	GACGGGATGACGACGACCAACCAACCATGCGACGACCAACCACTCCATCCACGACACGAC	526
QY	163	LeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThrAla	182
Db	527	CTGGCTACTCTGTCATGACTGCTGCGCGGAAAGGGCTCAGCCCACTGACTGCTGTGA	586
QY	183	LeuSerGlyAlaHisThrIleGlyGlnAlaArgGlyThrThrPheArgGlyArgIleTyr	202
Db	587	CTCACTGGAGCGCACACCGTGGGGTGCGGAGTGACCAACATTCGCGGTGCGGCTTAC	646
QY	203	GlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSer	222
Db	647	GSCGAGCTCAACATCAACGCGCGTTCGCGGCGTGTCCGGGGGAGACTCCGCAAGCC	706
QY	223	GlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAla	242
Db	707	GCGCGCGACACCAACCTGGCGCCCTGGAC---TCCACCCCAACGCTCTGCACAGCC	763

PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137528
PR	04-JUN-1999	9905-0137550
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138340
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139750
PR	21-JUN-1999	9905-0139763
PR	21-JUN-1999	9905-0139817
PR	22-JUN-1999	9905-0139899
PR	23-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	23-JUN-1999	9905-0140355
PR	24-JUN-1999	9905-0140694
PR	28-JUN-1999	9905-0140823
PR	28-JUN-1999	9905-0140891
PR	29-JUN-1999	9905-0140931
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141824
PR	01-JUL-1999	9905-0142154
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142820
PR	12-JUL-1999	9905-0142877
PR	13-JUL-1999	9905-0143324
PR	14-JUL-1999	9905-0143524
PR	15-JUL-1999	9905-0144005
PR	15-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	16-JUL-1999	9905-0144225
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	19-JUL-1999	9905-0144336
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144684
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0144815
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145148
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145276
PR	27-JUL-1999	9905-0145513
PR	27-JUL-1999	9905-0145518
PR	27-JUL-1999	9905-0145519
PR	28-JUL-1999	9905-0145521
PR	02-AUG-1999	9905-0146586
PR	02-AUG-1999	9905-0146588
PR	02-AUG-1999	9905-0146589
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302

PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149802.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:  
 Pred. No.: 1,08e-65 Length: 1218  
 Score: 870.50 Matches: 177  
 Percent Similarity: 71.29% Conservative: 44  
 Best Local Similarity: 57.10% Mismatches: 82  
 Query Match: 54.27% Indels: 7  
 DB: 21 Gaps: 3

US-10-047-825-4 (1-313) x AAC49694 (1-1218)

OY 10 LeuValAlaValSerLeuSerCysValAlaHisAlaGlnLeuSerProThrPheTyr 29  
 Db TTGATGATGATAAATATGATGTGGCCGTCACAGCTCCAGGCTCAGTGAACCGGATTTTAC 157

OY 30 AlaserSerCysProAsnLeuGlnSerIleValArgAlaAlaMetThrGlnAlaValAla 49  
 Db AAGCAAGCTGTCCATCATCTTCTCTTGTGTGACAGCAGTCCGAAACGGCCGTGGCC 217

OY 50 SerGlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAspCysPheValGln 69  
 Db AGAGAGCCTCGCATGGGCTTCTCTCTCTGTTGTTCATGATTTGTTCAT 277

OY 70 GlyCysAspGlySerIleLeuLeuAspAlaGly-----GlyGluLysThrAla 85  
 Db GGCTGTGACGATCTCTTACTGTGTGATGCACACCGTCTTTTGGAGGAAACCTCA 337

OY 86 GlyProAsnLeuAsnSerValArgGlyPheGluValIleAspThrIleLysArgAsnVal 105  
 Db GCACCCACCAATMACTCTGTGAGGGGTTCCAGATGATCCAAATCAACTTAAGCTT 397

OY 106 GluAlaAlaCysProGlyValAlaSerCysAlaAspIleLeuAlaValAlaAlaArgAsp 125  
 Db GAGAAATGTGCCCGGCGATCTCATGCCAGACATTTAGCCATCAGTCTCGGGAC 457

OY 126 GlyThrAsnLeuLeuGlyLysProThrTyrSerValProLeuGlyArgArgAspSerThr 145  
 Db TCCGTTCTCTCTAGTGAGCGGGGAGCGGTGAACCTTGGAAGAAGACTCTAGC 517

OY 146 ThrAlaSerAlaSerLeuAlaAsnSerAsnPro---ProProProThrAlaSerLeuGly 164  
 Db ACGGCGAAGCTTCGCGCGGCGCACTCCGAGTATCCCTCTCGGATCAGTACCTTACG 577

OY 165 ThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThrAlaLeuSer 184  
 Db AACCTCATAAACCGTTTCAAGCACAAGGTTTCCACAGTGACATGTCGCCCTCTCT 637

OY 185 GlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArgGlyArgIleTyrGlyAsp 204  
 Db GGTGCTCACACCATTTGACAGCCCAATGTGTACATTCAGAAACCGAATCTACACGCA 697

OY 205 ThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyGly 224  
 Db AGCAATATCGACACCTCTTTGGCCATCTCTAAGAGGAGAACTCTCTGCCACAGTGGC 757

OY 225 AspGlyAsn-----LeuAlaProIleAspValGlnThrProValArgPheAspThrAla 242  
 Db TCCGAGACACAAGAAAGCAATCTTACGTCGCTCCGATAGTTCGACACAGCGC 817

OY 243 TyrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnGluLeuPheAsn 262  
 Db TTCTACAGCACTTCTGAGCAAAAAGTTTGCTTACGTACGACCAAGCTCTTTAAT 877

OY 263 GlyIleSerGlnAspAlaLeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsnAla 282  
 Db AATGCTCTCCACCGCTCGCTCATATGATTACGCCAATCATGCTTCTTACCGC 937

OY 283 AspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGly 302  
 Db GACTTTCAAGGCAATGATTAAGTGGAGACATCAGCCCTCACCCGATCCTCAATGTT 997

OY 303 GlnIleArgArgAsnCysArgValAlaAsn 312  
 Db CAGATCCGCCAATAACTCTCGGAGGCCCAAC 1027

RESULT 12	ID	AAC36625	standard; DNA; 1220 BP.
XX	XX	AAC36625;	
XX	XX	17-OCT-2000 (first entry)	
XX	XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 14488.	
XX	XX	Hybridisation assay; genetic mapping; gene expression control;:	
XX	XX	protein identification; signal transduction pathway;	
XX	XX	metabolic pathway; promoter; termination sequence; ss.	
XX	XX	Arabidopsis thaliana.	
XX	XX	EP1033405-A2.	
XX	XX	06-SEP-2000.	
XX	XX	25-FEB-2000; 2000EP-0301439.	
XX	XX	25-FEB-1999; 99US-0121825.	
XX	XX	05-MAR-1999; 99US-0123180.	
XX	XX	09-MAR-1999; 99US-0123548.	
XX	XX	23-MAR-1999; 99US-0125788.	
XX	XX	25-MAR-1999; 99US-0126264.	
XX	XX	29-MAR-1999; 99US-0126785.	
XX	XX	01-APR-1999; 99US-0127462.	
XX	XX	06-APR-1999; 99US-0128234.	
XX	XX	08-APR-1999; 99US-0128714.	
XX	XX	16-APR-1999; 99US-0128645.	
XX	XX	19-APR-1999; 99US-0130077.	
XX	XX	21-APR-1999; 99US-0130449.	
XX	XX	23-APR-1999; 99US-0130510.	
XX	XX	23-APR-1999; 99US-0130891.	
XX	XX	28-APR-1999; 99US-0131449.	
XX	XX	30-APR-1999; 99US-0132048.	
XX	XX	30-APR-1999; 99US-0132407.	
XX	XX	04-MAY-1999; 99US-0132484.	
XX	XX	05-MAY-1999; 99US-0132485.	
XX	XX	06-MAY-1999; 99US-0132486.	
XX	XX	07-MAY-1999; 99US-0132487.	
XX	XX	11-MAY-1999; 99US-0132863.	
XX	XX	11-MAY-1999; 99US-0134256.	
XX	XX	14-MAY-1999; 99US-0134218.	
XX	XX	14-MAY-1999; 99US-0134219.	
XX	XX	14-MAY-1999; 99US-0134221.	
XX	XX	14-MAY-1999; 99US-0134370.	
XX	XX	18-MAY-1999; 99US-0134768.	
XX	XX	19-MAY-1999; 99US-0134941.	
XX	XX	20-MAY-1999; 99US-0135524.	
XX	XX	21-MAY-1999; 99US-0135523.	
XX	XX	24-MAY-1999; 99US-0135629.	
XX	XX	25-MAY-1999; 99US-0136021.	
XX	XX	27-MAY-1999; 99US-0136392.	
XX	XX	28-MAY-1999; 99US-0136782.	
XX	XX	01-JUN-1999; 99US-0137222.	
XX	XX	03-JUN-1999; 99US-0137528.	
XX	XX	04-JUN-1999; 99US-0137502.	
XX	XX	07-JUN-1999; 99US-0137724.	
XX	XX	08-JUN-1999; 99US-0138094.	
XX	XX	10-JUN-1999; 99US-0138540.	
XX	XX	10-JUN-1999; 99US-0138847.	
XX	XX	14-JUN-1999; 99US-0139119.	
XX	XX	16-JUN-1999; 99US-0139452.	
XX	XX	16-JUN-1999; 99US-0139453.	
XX	XX	17-JUN-1999; 99US-0139492.	
XX	XX	18-JUN-1999; 99US-0139454.	
XX	XX	18-JUN-1999; 99US-0139455.	
XX	XX	18-JUN-1999; 99US-0139456.	
XX	XX	18-JUN-1999; 99US-0139457.	

PR	18-JUN-1999	99US-01394.58
PR	18-JUN-1999	99US-01394.58
PR	18-JUN-1999	99US-01394.60
PR	18-JUN-1999	99US-01394.61
PR	18-JUN-1999	99US-01394.62
PR	18-JUN-1999	99US-01394.63
PR	18-JUN-1999	99US-01394.65
PR	18-JUN-1999	99US-01396.73
PR	21-JUN-1999	99US-01396.17
PR	22-JUN-1999	99US-01396.17
PR	23-JUN-1999	99US-01403.53
PR	24-JUN-1999	99US-01403.54
PR	24-JUN-1999	99US-01406.93
PR	28-JUN-1999	99US-01408.25
PR	29-JUN-1999	99US-01409.91
PR	30-JUN-1999	99US-01412.87
PR	14-JUL-1999	99US-01435.42
PR	15-JUL-1999	99US-01436.24
PR	15-JUL-1999	99US-01440.05
PR	15-JUL-1999	99US-01440.85
PR	15-JUL-1999	99US-01440.86
PR	15-JUL-1999	99US-01443.25
PR	15-JUL-1999	99US-01443.31
PR	15-JUL-1999	99US-01443.32
PR	15-JUL-1999	99US-01443.33
PR	15-JUL-1999	99US-01443.37
PR	13-JUL-1999	99US-01443.34
PR	19-JUL-1999	99US-01443.35
PR	20-JUL-1999	99US-01443.52
PR	20-JUL-1999	99US-01446.32
PR	20-JUL-1999	99US-01446.84
PR	21-JUL-1999	99US-01448.14
PR	21-JUL-1999	99US-01450.86
PR	21-JUL-1999	99US-01450.88
PR	22-JUL-1999	99US-01450.87
PR	22-JUL-1999	99US-01450.87
PR	22-JUL-1999	99US-01451.92
PR	23-JUL-1999	99US-01451.45
PR	23-JUL-1999	99US-01452.18
PR	23-JUL-1999	99US-01452.24
PR	26-JUL-1999	99US-01453.76
PR	27-JUL-1999	99US-01453.13
PR	27-JUL-1999	99US-01455.18
PR	27-JUL-1999	99US-01455.19
PR	28-JUL-1999	99US-01455.1
PR	02-AUG-1999	99US-01463.86
PR	02-AUG-1999	99US-01463.88
PR	02-AUG-1999	99US-01465.88
PR	03-AUG-1999	99US-01470.38
PR	04-AUG-1999	99US-01472.04
PR	04-AUG-1999	99US-01473.02
PR	05-AUG-1999	99US-01471.92
PR	05-AUG-1999	99US-01472.60
PR	06-AUG-1999	99US-01473.03
PR	06-AUG-1999	99US-01474.16
PR	09-AUG-1999	99US-01474.93
PR	09-AUG-1999	99US-01476.35
PR	10-AUG-1999	99US-01481.71
PR	11-AUG-1999	99US-01483.19
PR	12-AUG-1999	99US-01483.41
PR	13-AUG-1999	99US-01485.65
PR	13-AUG-1999	99US-01486.84
PR	17-AUG-1999	99US-01493.68
PR	17-AUG-1999	99US-01491.75
PR	18-AUG-1999	99US-01494.26
PR	20-AUG-1999	99US-01497.22

```

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.

```

## Alignment Scores:

```

Pred. No.: 1 966-65
Score: 867.50
Percent Similarity: 71.29%
Best Local Similarity: 56.77%
Query Match: 54.08%
DB: 21

```

US-10-047-825-4 (1-313) x AAC36625 (1-1220)

```

QY 10 leuValAlaValserLeuSerCysValAlaHisAlaGlnLeuSerProThrPheTyr 29
DB 100 TTGATGTGATGATGATGTTGGCTGCACAGTCGAGCTCAGTGAACCGTACTTTTAC 159
QY 30 AlaserSerCysProAsnLeuGlnSerIleValArgAlaIleMetThrGlnAlaValAla 49

```

```

DB 160 AAGAAAGCTGTCCATCATTTGCTTGTGTCGAGACGAGTGAACGGGGCCGTGGCC 219
QY 50 SerGlnGlnArgMetGlyAlaSerIleLeuArgLeuPhePheHisAspCysPheValGln 69
DB 220 AGAAGGCTCCGATGGGATGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 279
QY 70 GlyCysAspGlySerIleLeuLeuAspAlaGly-----GlyGlnLysThrAla 85
DB 280 GGGGTGACGAGATCTTACTGTGGATGACACACCGCTCTTTTGGAGAGAAAACCTCA 339
QY 86 GlyProAsnLeuAsnSerValArgGlyPheGlnValIleAspThrIleLysArgAsnVal 105
DB 340 GGACCCAGCAATTAAGCTGTGAGGGGTTCGAGATGATGACAAATTAAGTTTAAGTT 399
QY 106 GluAlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAsp 125
DB 400 GAGAAATGTGCGCGGCGCATCGTCTCATGCGAGACATTCAGCCATCATCGTCTGGGAC 459
QY 126 GlyThrAsnLeuLeuGlyGlyProThrTTPSerValProLeuGlyArgArgAspSerThr 145
DB 460 TCCGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 519
QY 146 ThrAlaSerAlaSerIleAlaAsnSerAsnPro---ProProProThrAlaSerLeuGly 164
DB 520 ACGGCGAACTTCGGCGCGCGGAGCTCGGAGTCATCCCTCCGATCATCACTTACG 579
QY 165 ThrIleuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThrAlaLeuSer 184
DB 580 AACCTCATTAACCGTTCAAGCAAGGTTGTCACACGAGATGTCGACATGTCGCTCTCT 639
QY 185 GlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArgGlyArgIleTyrGlyAsp 204
DB 640 GGTTCACACCATGATGAGCAGCCCATGTTACTTCAGAAACGAAATCATCAACGCA 699
QY 205 ThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyGly 224
DB 700 AGCATATCGACACCTTCTGCGCATCTCTAAACGGGAACTGTCGCGACACAGTGGC 759
QY 225 AspGlyAsn-----LeuAlaProIleAspValGlnThrProValArgPheAspThrAla 242
DB 760 TCCGAGAGCAACAGAAAGCCATCTTGACGTCGCTCCGATGAGTGCACACGCGC 819
QY 243 TyrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnGlnLeuPheAsn 262
DB 820 TTCTACAGCAACTCTGACCAAAAAGGTTGCTTACGTACGACCAAGTCTCTTTAAT 879
QY 263 GlyGlySerGlnAspAlaLeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsnAla 282
DB 880 AATGCTCTACCGCATCGTCTGATAGCTACAGCCACATTCGATCCCTTACCGC 939
QY 283 AspheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGly 302
DB 940 GACTTTCAGAGGCAATGATTAAGATGAGACATCAGCCCTTCACCGATCAATGCT 999
QY 303 GlnIleArgArgAsnCysArgValValAsn 312
DB 1000 CAGATCCGCCAAACCTGCGAGGCCCAAC 1029

```

## RESULT 13

```

AA68192
ID AA68192 standard; DNA; 801 BP.
AC AA68192;
XX
XX 24-OCT-2000 (first entry)
XX
XX Pinus radiata peroxidase nucleotide sequence SEQ ID NO:368.
XX
XX Plant; Lignin; Lignin biosynthetic pathway; Eucalyptus grandis;
XX
XX Pinus radiata; Monterey pine; ds.
XX
XX Pinus radiata.
OS

```



XX PN WO200022099-A1.  
 XX PD 20-APR-2000.  
 XX PF 06-OCT-1999; 99WO-NZ00168.  
 XX PR 09-OCT-1998; 98US-0169789.  
 XX PR 14-JUL-1999; 99US-0143811.  
 XX PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX PI Bloksberg LN, Havukala IJ;  
 XX DR WPI: 2000-317962/27.  
 XX PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic  
 PT pathway useful for producing transgenic plants especially eucalyptus  
 PT and pine species having altered lignin content, composition and  
 PT structure  
 XX PS Claim 1, Page 189; 213pp; English.

CC The present invention describes isolated polynucleotides and proteins  
 CC encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),  
 CC coumarate 3-hydroxylase (C3H), phenolase (PML), O-methyl transferase  
 CC (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase  
 CC (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),  
 CC coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),  
 CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,  
 CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,  
 CC coumarate CoA ligase, cytochrome P450 1XX1A, diaphenol oxidase, flavanol  
 CC glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,  
 CC which are involved in the lignin biosynthetic pathway. The  
 CC polynucleotides can be used for modulating lignin content, lignin  
 CC composition and the structure of a plant, especially eucalyptus and pine  
 CC species, and for modifying the activity of an enzyme involved in lignin  
 CC biosynthetic pathway, and for producing a plant having altered lignin  
 CC content, composition and structure. They can be used for designing probes  
 CC and primers useful for detecting similar DNA and RNA sequences in any  
 CC organism and for PCR amplification. The lignin content can be efficiently  
 CC modified using the polynucleotides. AAG67908 to AAG68201 and AAB16341 to  
 CC AAB16449 represent polynucleotide and protein sequences used in the  
 CC exemplification of the present invention.

CC XX Sequence 801 BP; 195 A; 217 C; 195 G; 194 T; 0 other;

#### Alignment Scores:

Pred. No.: 8,15e-64 Length: 801  
 Score: 846.00 Matches: 170  
 Percent Similarity: 75.94% Conservative: 32  
 Best Local Similarity: 63.91% Mismatches: 58  
 Query Match: 52.74% Indels: 6  
 DB: 21 Gaps: 2

US-10-047-825-4 (1-313) x AAG68192 (1-801)

QY 1 MetAlaSerProThrLeuMetGlnCysLeuValAlaValSerLeuLeuSerCys----- 18  
 DB 4 ATGGCTTCGTTACAGCAATGCGATCTGCGCTTTATGATGATGTTGCGACC 63  
 QY 19 ValAlaHisAlaGlnLeuSerProThrPheTyrAlaSerSerCysProAsnLeuGlnSer 38  
 DB 64 GTTGGGACGGCGCAAGCTTTATATATACATATGTCCTCCAAACTACTCTCA 123  
 QY 39 IleValArgAlaAlaMetThrGlnAlaValAlaSerGlnArgMetGlyAlaSerLeu 58  
 DB 124 ACGGTGACGGCGCTGTGAACAGCGGTGCGCAAGAGAGCCATGCGGGGCTCCTC 183  
 QY 59 LeuAlaGlnPhePheHisArgCysPheValGlnGlyCysAspGlySerIleLeuLeuAsp 78  
 DB 184 CTCGCGCTTCACTTCAAGACTGCTTGTATGATGATGATGATGATGATGATGATG 243

QY 79 AlaGly-----GlyGluLeuThrAlaGlyProAsnLeuAsnSerValAlaGly 94  
 DB 244 GACTCTTCGACTCTAAGTGGAGAGAACCCCGCTTCCCAACAATTCGGCAAGGGG 303  
 QY 95 PheGluValIleAspThrIleLysArgAsnValGluAlaAlaLysProGlyValValSer 114  
 DB 304 TTCATGTGATAGACACCATCAAGTCTCAAGTGAACACACTTGCACAGTGGTGTGCG 363  
 QY 115 CysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuGlnGlyProThr 134  
 DB 364 TGGCAGATATTTGGCTATTGCTGAGATTCCTGTTGTCAGTTGGAGGCTCCACA 423  
 QY 135 TrpSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSer 154  
 DB 424 TGGACAGTGCAGCTGGGAGGAGAGACTCCAGACTGCGACGCTAAGTGGTGAACAC 483  
 QY 155 AsnProProProProThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGly 174  
 DB 484 AACATTTCCGGCTCTACTTCAATCTCAGTCTCATCTCATTTATTTCAAGCTCAGGGT 543  
 QY 175 LeuSerProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCys 194  
 DB 544 CTTCACAGAGAGACATGTTGCTCTATCAGTGGCAGACACCATAGGCCAAGCGGGTGC 603  
 QY 195 ThrThrPheArgGlyArgIleTyrGlyAspThrAspIleAsnAlaSerPheAlaAlaLeu 214  
 DB 604 ACAACTTCAGGGCGCGCATCTACACCAATCAATTAATGAGCATGCAACTTCC 663  
 QY 215 ArgGlnGlnThrCysProArgSerGlyLysPheGlyAsnLeuAlaProIleAspValGln 234  
 DB 664 CTGAAGACAACTGTCCAGTACAGAGCGCAACAACCTGTGCTCCATTGGATGCTGT 723  
 QY 235 ThrProValArgPheAspThrAlaTyrPheThrAsnLeuLeuSerArgArgGlyLeuPhe 254  
 DB 724 ACTCCACACTACTTTGACATCACTCACTCACTCACTCACTCACTCACTCACTCTC 783  
 QY 255 HisSerAspGlnLeu 260  
 DB 784 CACTCCGACACAGACACTG 801

RESULT 14  
 AAG63720  
 ID AAG63720 standard; cDNA; 1555 BP.  
 AC XX AAG63720;  
 DT 12-APR-1999 (first entry)  
 XX XX  
 DE Malze per5 root preferential cationic peroxidase cDNA.  
 XX XX  
 KW peroxidase; per5 gene; malze; corn; transgenic plant; promoter;  
 XX root; ss.  
 OS Zea mays.  
 FH Key Location/Qualifiers  
 FT CDS 53..1054  
 FT sig\_peptide /\*tag= a  
 FT mat\_peptide /\*tag= b  
 FT polyA\_signal /\*tag= c  
 FT /\*tag= d  
 PN WO9856921-A1.  
 PD 17-DEC-1998.  
 PF 10-JUN-1998; 98WO-US11921.  
 PR 12-JUN-1997; 97US-0049752.







GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 3, 2003, 12:15:18 ; Search time 46 Seconds  
(without alignments)  
2086.735 Million cell updates/sec

Title: US-10-047-825-4  
Perfect score: 1604  
Sequence: 1 MASPTMOCLVAIVSLSCVA.....VGLTGTAGQIRRCRVVNS 313

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US10047825/rnat.26042003.111959.7157/app.query.fasta\_1.455  
-DB=Issued\_Patents\_NA\_QPMT-fastbp -SUFFIX=tnl -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US10047825.@cgn2.1.32.@rnat.26042003.111959.7157 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMBOUR=120  
-WARN\_TIMBOUR=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA.\*

- 1: /cgn2.6/ptodata/1/lna/5A.COMB.seq.\*
- 2: /cgn2.6/ptodata/1/lna/5B.COMB.seq.\*
- 3: /cgn2.6/ptodata/1/lna/6A.COMB.seq.\*
- 4: /cgn2.6/ptodata/1/lna/6B.COMB.seq.\*
- 5: /cgn2.6/ptodata/1/lna/PCTUS.COMB.seq.\*
- 6: /cgn2.6/ptodata/1/lna/Backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	967	60.3	1224	4	US-09-615-192A-363 Sequence 363, App
2	950	59.2	1171	4	US-09-615-192A-369 Sequence 369, App
3	899.5	56.1	1391	4	US-09-615-192A-352 Sequence 352, App
4	846	52.7	801	4	US-09-615-192A-368 Sequence 368, App
5	832	51.2	1354	4	US-09-097-319A-5 Sequence 5, Appl
6	807	50.3	916	4	US-09-615-192A-361 Sequence 361, App
7	723	45.1	975	4	US-09-365-150-3 Sequence 3, Appl
8	721.5	45.0	1270	1	US-08-378-588-23 Sequence 23, Appl
9	721.5	45.0	1270	5	PCT-US94-11121-23 Sequence 23, Appl
10	721.5	44.5	975	4	US-09-365-150-1 Sequence 1, Appl
11	714.5	44.5	975	4	US-09-365-150-4 Sequence 4, Appl
12	714.5	44.5	975	4	US-09-365-150-4 Sequence 4, Appl

13	664	41.4	1522	4	US-09-615-192A-371 Sequence 371, App
14	628	39.2	6550	4	US-09-097-319A-1 Sequence 1, Appl
15	624	38.9	646	4	US-09-615-192A-365 Sequence 365, App
16	622.5	38.8	1131	1	US-08-181-271A-22 Sequence 22, App
17	622.5	38.8	1131	1	US-08-449-315-22 Sequence 22, App
18	622.5	38.8	1131	1	US-08-444-803-22 Sequence 22, App
19	622.5	38.8	1131	1	US-08-449-043-22 Sequence 22, App
20	622.5	38.8	1131	1	US-08-456-265A-22 Sequence 22, App
21	622.5	38.8	1131	1	US-08-455-416-22 Sequence 22, App
22	622.5	38.8	1131	1	US-08-455-244-22 Sequence 22, App
23	622.5	38.8	1131	1	US-08-454-876-22 Sequence 22, App
24	622.5	38.8	1131	2	US-08-457-364-22 Sequence 22, App
25	622.5	38.8	1131	2	US-08-456-262-22 Sequence 22, App
26	622.5	38.8	1131	2	US-08-456-240-22 Sequence 22, App
27	622.5	38.8	1131	2	US-08-455-736-22 Sequence 22, App
28	622.5	38.8	1131	2	US-08-971-217-22 Sequence 22, App
29	622.5	38.8	1131	2	US-09-350-600-22 Sequence 22, App
30	605.5	37.7	1326	2	US-08-671-320-12 Sequence 12, App
31	605.5	37.7	1326	2	US-08-868-577-12 Sequence 12, App
32	604.5	37.7	1314	2	US-08-868-577-10 Sequence 10, App
33	604.5	37.7	1315	2	US-08-671-320-10 Sequence 10, App
34	585.5	36.5	1072	1	US-08-190-029A-9 Sequence 9, Appl
35	585.5	36.5	1072	2	US-08-462-695-9 Sequence 9, Appl
36	573	35.7	1064	1	US-08-378-588-15 Sequence 15, App
37	573	35.7	1064	2	US-08-811-094-15 Sequence 15, App
38	573	35.7	1064	2	PCT-US94-11121-15 Sequence 15, App
39	551.5	34.4	1191	2	US-08-671-320-14 Sequence 14, App
40	551.5	34.4	1191	2	US-08-868-577-14 Sequence 14, App
41	547.5	34.1	1167	2	US-08-671-320-16 Sequence 16, App
42	547.5	34.1	1167	2	US-08-868-577-16 Sequence 16, App
43	542	33.8	1181	1	US-08-181-271A-27 Sequence 27, Appl
44	542	33.8	1181	1	US-08-449-315-27 Sequence 27, Appl
45	542	33.8	1181	1	US-08-444-803-27 Sequence 27, Appl

## ALIGNMENTS

RESULT 1  
US-09-615-192A-363  
Sequence 363, Application US/09615192A  
Patent No. 6410718  
GENERAL INFORMATION:  
APPLICANT: Blockberg, Leonard N.  
TITLE OF INVENTION: Materials and Methods for the  
FILE REFERENCE: 11000.1003c4u  
CURRENT FILING DATE: US/09/615,192A  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 08/975,316  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11  
PRIOR APPLICATION NUMBER: US 09/169,789  
PRIOR FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 405  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 363  
LENGTH: 1224  
TYPE: DNA  
ORGANISM: Pinus radiata  
US-09-615-192A-363  
Alignment Scores:  
Pred. No.: 1.9e-91  
Score: 967.00  
Percent Similarity: 72.41%  
Best Local Similarity: 59.87%  
Query Match: 60.29%  
DB: 4  
Gaps: 12  
US-10-047-825-4 (1-313) x US-09-615-192A-363 (1-1224)

```

Oy      6 LeuMeGlnCysLeuValAlaValSerIleuSerCys----- 18
Db      56 ATGATGAGGACTCTATATGTCATGGCTTAATGGCTGTGTTAGCCTTCATCATATA 11
Oy      19 ---ValAlaHisIAsGInIleuSerProThrPheIyAlaSerSerCysProAsnIeuGIn 37
Db      116 AACGCGTGAATGGCAGGTGAGCTCAACGTTTTATGCGCAAAATCGTGTCCGAGGTTGCCA 17
Oy      38 SerIleValAlaArgAlaAlaMetThrGlnAlaValaISerGInIleArgMetGlyAlaSer 57
Db      176 TCATATGTGAATCACTGGTGAAGCAACGGTACTTAAGAGAAAGAAAGGAGGCGTCC 23
Oy      58 LeuIleuArgIleuPhePheHisAspCysPheValGlnIleCysAspGlySerIleIleuLeu 77
Db      236 TTGGTCCGCGCTTCACATTCACGATTCGTCGTCAACGGGATGGCATGGTTCATCTATTG 29
Oy      78 AspAlaIy-----GlyIleuIyThrAlaGlyProAsnIleuAsnSerValArg 93
Db      296 GATGACAAACGCTACGTTTACCGGAGAAAAGCTCAGGCCCAAAACGGCAATTCGTGGCA 35
Oy      94 GlyPheGluValIleAspThrIleIyAsnArgAsnValGlnAlaIaCysProGlyValAla 11
Db      356 GCGCTCGAGTAATTACAGCAATTAAACTCAATGGAGGCACGCTTCAGTGGAGTCGG 41
Oy      114 SerCysAlaAspIleIleuAlaIleuAlaArgAspGlyThrAsnIleuIleuGlyIlePro 13
Db      416 TCGGTGTCAACACTCTCCACACTTCGCTCGTCACTCTATGTCGTAACCTCAAGGCCCA 47
Oy      134 ThrTrpSerAlaProIleuIyArgArgAspSerThrThrIleuIleSerIleuPheGlyArgGIn 15
Db      476 ACATGACGGTAATGCTTGGAAGGCGAAGCTCCACGACTCGGAGATTAAACGCTGCAAAAC 53
Oy      154 SerAsnProProProThrIleAspIleuGlnIyThrIleuIleSerIleuPheGlyArgGIn 17
Db      536 AACAACTTCATCTCCGCTTCGATCGAGACACTCATCTCATCTTTCAAGCTCAC 59
Oy      174 GlyIleuSerProArgAspMetThrAlaIleuSerGlyAlaHisThrIleGlyGlnAlaArg 19
Db      596 GCGTTCCTTCAACAAACCTTTCACCTCTCAGGTGCTCATAAATGGTCCATATACGA 65
Oy      194 CysThrThrPheArgGlyIyArgIleTyGlnIyAspThrAspIleAsnAlaSerPheAlaIa 21
Db      656 TGGCGCTTTTTCAGACTGGATCTACAAACGAAGAAACATTAACGCTGCTTCGCTACA 71
Oy      214 LeuArgGlnGInThrCysProArgSerGlyGlyAspGlyIleuAlaProIleAspVal 23
Db      716 TCTGTAAAGCCAAACGCCCCAGCGCTGTGGCGACAGCAACCTCTCTCCTTAGATGGC 77
Oy      234 GlnThrProValAlaArgPheAspThrAlaIyPheThrAsnIleuSerArgArgGlyIleu 25
Db      776 GTCACTTCATTCACATTATACAAACAAGTATTACTTATCTTAATAATACGAAGAAGCTT 83
Oy      254 PheHisSerAspGlnIleuIleuPheAsnGlyIySerGlnAspAlaIleuValArgGlnIy 27
Db      836 CTCGACTCGACACGAGCGCTTTAATGAGGATTCACAGATTCCTCAGTTACTGCGTAC 89
Oy      274 SerAlaSerAlaSerIleuPheAsnAlaAspPheValAlaAlaMetIleArgMetGlyAsn 29
Db      896 AGCAGCAATTCAGAACGCTTTCTTTAATACCTTTACACTGCGACGTGGAAGATGGGAAT 95
Oy      294 ValGlyValIleuThrGlyThrAlaGlyGlnIleIyArgArgAsnCysArgValAlaAsn 312
Db      956 ATTAGCCCTCTCAGTGGCACTAAGCGGCAAAATCCGCAAAATCGCAAGAAAGTCAAT 1012

RESULT_2
US-09-615-192A-369
: Sequence 369, Application us/09615192A
: Patent No. 6410718
: GENERAL INFORMATION:
: APPLICANT: Bloksberg, Leonard N.
: APPLICANT: Havukala, Ilka
: TITLE OF INVENTION: Materials and Methods for the
: TITLE OF INVENTION: Modification of Plant Lignin Content

```

```

? FILE REFERENCE: 11000.1003040U
? CURRENT APPLICATION NUMBER: US/09/615.192A
? CURRENT FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 08/975.316
? PRIOR FILING DATE: 1997-11-21
? PRIOR APPLICATION NUMBER: US 08/713.000
? PRIOR FILING DATE: 1996-09-11
? PRIOR APPLICATION NUMBER: US 09/169.789
? PRIOR FILING DATE: 1998-10-09
? NUMBER OF SEQ ID NOS: 405
? SOFTWARE: FASTSeq for Windows Version 3.0
? SEQ ID NO 369
? LENGTH: 1171
? TYPE: DNA
? ORGANISM: Plinus radiata
US-09-615-192A-369

Alignment Scores:
Pred. No.: 1,05e-89 Length: 1171
Score: 950.00 Matches: 186
Percent Similarity: 73.25% Conservative: 44
Best Local Similarity: 59.24% Mismatches: 74
Query Match: 59.23% Indels: 10
DB: Gaps: 3

US-10-047-825-4 (1-313) x US-09-615-192A-369 (1-1171)
?
QY 9 CysleuValAlaIalSerLeuLeuSerCys-----ValAlaHisAlaGlnLeu 24
   |||||: ||: ||: |||
Db 70 TGTATAGCTGTAATGGTGTATTATCGTGTCTTATACAGTACTGATGGACGACTG 129
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 25 SerProthrPheTyraIaSerSerCysProAsnLeuGlnSerIleValArgAlaIaMet 44
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 ACCTCCAGCTTTACGAAATCTTGCCCGACGACCTGTCGTAGTGAAGGCCGCACTG 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 45 ThrGlnAlaValAlaIalSerGlnArgMetGlyValaSerLeuLeuArgLeuPhePheHis 64
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 AAGCAAGGGGTCGCTAACGAGAAAGGATGGGTGGCTTTGCTTCCGCCCTTCAC 249
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 65 AspCysPheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGly- - - - - 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 GACTGCTTCGTTAATGATGGATGGATGGTCGCTTCGTTGACGATTCCTGCACCTTACT 309
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 81 GlyGlyIysThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlyValIleAspThr 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 GCGGGAAGACAGCTAATCCAAATCCAAATTCGTCGAGAGGATTCGACGTAAAGATACC 368
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 101 IleValArgAsnValGlnAlaAlaCysProGlyValaSerCysAlaAspIleLeuAla 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 ATTAAGACCAATGTCGAAAGACCTTGACGTGGAGTCGTTCTCTGCAGACATTCGCC 429
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 LeuAlaIalArgAspGlyThrAsnLeuGlnGlyProThrPseValProLeuGly 140
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 ATTTGCTGCTCGATTCGTTGTTGAATCGGGCGCTCCCTTCATGACAGTAATCTTGGA 489
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 141 ArgArgAspSerThrThrAlaSerAlaSerLeuIlaAsnSerAsnProProProThr 166
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 AAGCGAGACTGCACACAGCTTAGAAAGCGGTGCAACCAATATATCCGCTCCGACT 549
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 161 AlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMet 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 TCCACTTCGAGAACCTCATCTCACTATATCCAAACGCGGAGCTCTCCGAAAGAAATG 609
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 ThrAlaLeuSerGlyAlaHisThrIleGlyAlaIalArgCysThrTrpPheArgGlyArg 200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 GTTGCACCTTTCGGGGTCATATACATCCGGGAGGCGGATTCGAAGATTTTCAGAGCCAT 669
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 201 IleTyrgIlyAspThrAspIleAsnAlaSerPheAlaIalLeuArgGlnGlnThrCysPro 220
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 670 ATTTCACAGAGACCAATAGACATGACGTCCGTAACGCCCACTTCATGCTTCAAGTGTCCG 729
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 221 ArgSer-----GlyGlyAspGlyAsnLeuAlaProIleAspValGlnTrpProValArg 238
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 730 AGTACCACAGGCTCCGAGACAGCACTTGTGCGCATTTGATTATATGATCCACTGNG 789  
QY 239 PheASPThrAlatyrPheRhsLeuLeuSerArgArgLysLeuPheHisSerAspGln 258  
Db 790 TTTGACAAAACCTATTACGCGACCTGAAAGCAAAAGAGCTTCTCCACTCCGACG 849  
QY 259 GluLeuPheAsnGlySerGlnAspAlaLeuValArgGlnTyrSerAlaSerAlaSer 278  
Db 850 GAACCTTCACAGGAGGCTCCACTGATTCACAGGTACAGTACGCTCCACACAGAAC 909  
QY 279 LeuPheAsnAlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThr 298  
Db 910 ACCCTTCTTCGATTTCTGCGGACATGTTAGATGGAATAATCAACCTCTTACC 969  
QY 299 GlyThrAlaGlyGlnIleArgArgAsnGlyValAlaValAsn 312  
Db 970 GGCACCAAGCGACAGATCCCAAGAACTCCAGAAAGCCCAAC 1011

RESULT 3  
US-09-615-192A-352  
; Sequence 352, Application US/09615192A  
; Patent No. 6410718  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; APPLICANT: Havukkala, Ilkka  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1003c4U  
; CURRENT APPLICATION NUMBER: US/09/615,192A  
; CURRENT FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 08/975,316  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: US 08/713,000  
; PRIOR FILING DATE: 1996-09-11  
; PRIOR APPLICATION NUMBER: US 09/169,789  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 405  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 352  
; LENGTH: 1391  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-09-615-192A-352

Alignment Scores:  
Pred. No.: 2,47e-84 Length: 1391  
Score: 899.50 Matches: 180  
Best Similarity: 71.528 Conservative: 46  
Best Local Similarity: 56.968 Mismatches: 85  
Query Match: 56.088 Indels: 5  
DB: 4 Gaps: 2

US-10-047-825-4 (1-313) x US-09-615-192A-352 (1-1391)

QY 1 MetaLaserProThrLeuMetGlnCysLeuValAlaValSerLeuLeuSerCysValAla 20  
Db 246 ATGTGCGCTCTCCGCTCCGCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 302  
QY 21 HisAlaGlnLeuSerProThrPheTyrAlaSerSerCysProAsnLeuGlnSerIleVal 40  
Db 303 CGTTGATGCTCTCCGCTCCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 362  
QY 41 ArgAlaAlaMetThrGlnAlaValAlaValSerGlnArgMetGlyAlaSerLeuLeuArg 60  
Db 363 CGAGCTGGAATACCAAGAAATCAAGAAATGACCCGAGCGCGCTGCTGCTGCTGCTGCTG 422  
QY 61 LeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGly 80  
Db 423 CTGCACCTTCATGACTCTCCGCTCAATGGGTGCGATGCGATGATGTTGTTGATGACAGC 482  
QY 81 -----GlyGluTyrThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlu 96  
Db 483 CCTAGCTTCTGCTGGCGAGAAACAGACAGCTCCAGACACACATTCCTGAGAGGCTTCGAA 542

QY 97 ValIleAspThrIleTyrArgAsnValGlnAlaAlaCysProGlyValValSerCysAla 116  
Db 543 GTGATCGACCCGATCAGAGCTAGTCTGAGAGAGAGTCCCTGAGAGTCTTCTGTCGA 602  
QY 117 AspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGlyProThrTyrSer 136  
Db 603 GATATCGTTCGCTGCGGCTGCTGCGACTCACTGCTTCATTTGAGAGGCTTCATGAGAC 662  
QY 137 ValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAspPro 156  
Db 663 GTAACCTTAGGAGAAAGATTCATTAATCTGACAGGAGACCTTGTCAACCTTCATA 722  
QY 157 PropProProThrAlaSerGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSer 176  
Db 723 CCTCCACTTCTTAATCTCAGTCTCTCATTAACAGCTTCGCTGCTGAGGCTTCCTCA 782  
QY 177 ProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyAlaAlaArgCysThrThr 196  
Db 783 GTCAAGAACATGTTGCTCTTCTGTTACATACATTCATTCCTAGGAGATGCACTTCC 842  
QY 197 PheArgGlyArgGlyIleTyrGlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGln 216  
Db 843 TTCGAGAGACGATCTACACGACTCGAACATAGATACATCTTCGCTCAATTAATTCAG 902  
QY 217 GlnThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrPro 236  
Db 903 GAAGATCTCCAGAGATTGGAATGATGCTGCTTCAAGAGCTAGACATCCAAAGCCG 962  
QY 237 ValArgPheAspThrAlaTyrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSer 256  
Db 963 ACCCTCTTGAACAACCTTACTACACAAATTTACTGCAAGAAAGGCGCTTCTTCACTCT 1022  
QY 257 AspGlnGluLeuPheAsnGlyGlySerGlnAspAlaLeuValArgGlnTyrSerAlaSer 276  
Db 1023 GATCAGAGCTCTTAATGAGAGCTCTGCTGATTCAGTCAAGAAATGATGCTGCGAC 1082  
QY 277 AlaSerLeuPheAsnAlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyVal 296  
Db 1083 ACAGCAAAATTTTCCGAGATTTTCCAGCAATGATTCAAAGCAATGAGAGAAATTAAGCCC 1142  
QY 297 LeuThrGlyThrAlaGlyGlnIleArgArgAsnGlyValValAsn 312  
Db 1143 CCCAAAGAGACAAATGCTCAATTAAGGAAATTAAGGAAAGTGAAC 1190

RESULT 4  
US-09-615-192A-368  
; Sequence 368, Application US/09615192A  
; Patent No. 6410718  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; APPLICANT: Havukkala, Ilkka  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1003c4U  
; CURRENT APPLICATION NUMBER: US/09/615,192A  
; CURRENT FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 08/975,316  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: US 08/713,000  
; PRIOR FILING DATE: 1996-09-11  
; PRIOR APPLICATION NUMBER: US 09/169,789  
; NUMBER OF SEQ ID NOS: 405  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 368  
; LENGTH: 801  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-615-192A-368

Alignment Scores:  
Pred. No.: 4e-79 Length: 801

Score: 846.00 Matches: 170  
 Percent Similarity: 75.94% Conservative: 32  
 Best Local Similarity: 63.91% Mismatches: 58  
 Query Match: 52.74% Indels: 6  
 DB: 4 Gaps: 2

US-10-047-825-4 (1-313) x US-09-615-192A-368 (1-801)

```

QY 1 MetalSerProThreumetGlnCysLeuValAlaValSerLeuSerCys----- 18
   ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 ATGGCTTCGTTTACACAAATGCGATCTCGGCTTATGCTTGTGTGATGTTGCGACC 63
QY 19 ValAlaHisAlaGlnLeuSerProThrPheTyrAlaSerSerCysProAsnLeuGlnSer 38
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GTTGCGTACGCGCAGCTTATGCGCAATTTTATATACATCAATGCTCCAACTACTCTCA 123
QY 39 IleValArgAlaAlaMetThrGlnAlaValAlaSerGlnArgMetGlyAlaSerLeu 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 ACCGTACAGCGCGCTGTGAAGCAAGCGGTGGCCACAGACAGCGATGGGGCGCTC 183
QY 59 LeuArgLeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeuAsp 78
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 CTCGGCTTCACCTTCACACACTGCTCGTATGTTGATGGTGGATGGGTGCTGCTGAC 243
QY 79 AlaGly-----GlyGluYstIrrAlaGlyProAsnLeuAsnSerValArgGly 94
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 GACTCTTCGACTCTTAAGTGGAGAGACCGCGTCCCAACAAATCGCGCAAGGGT 303
QY 95 PheGluValIleAspThrIleYstArgAsnValGluAlaAlaCysProGlyValAlaSer 114
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 TTGGATGTATGACACCAATCAAGTCTCAAGTGAAGCGATTGCGAGTGAAGTGTGCG 363
QY 115 CysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuGlyGlyProThr 134
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 364 TCGGAGATATTTGGCTATGCTGTGCTAGATTCGTTGCTGCTGAGTGGAGCCCAACA 423
QY 135 ThrSerValProLeuGlyArgArgAspSerThrAlaSerAlaSerLeuAlaAsnSer 154
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 TGGACAGTCAGCTGGGAGAGAGACTCCAGACTGCGCAAGTGTGCTCAACAAC 483
QY 155 AspProProProProThrAlaSerLeuGlyThrIleSerLeuPheGlyArgGlnGly 174
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 484 AACATTCGGCTCTACTCTTATCTCACTGCTCTCATCTCTTATTTTCAAGCTTAAGGT 543
QY 175 LeuSerProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyAlaArgCys 194
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 544 CTTTCCACAGAGACATGCTGTGCTATGAGGTGGCAGACCATAGCCCAAGCGGCTGC 603
QY 195 ThrThrPheArgGlyArgGlyIleTyrGlyAspThrAspIleAsnAlaSerPheAlaLeu 214
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 ACAAGCTTCAAGGCGCGCATCTACAAAGCAATCAACATTATGACGCAATGCAACTTCC 663
QY 215 ArgGlnGlnThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGln 234
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 664 CTGAACAACAACGCTCGGACTACAGAACGCAACACACTCTCACATGATGATGCTGT 723
QY 235 ThrProValArgPheAspThrAlaTyrPheThrAsnLeuSerArgArgGlyLeuPhe 254
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 724 ACTCCACTACGCTTGTGACATCACTACTCAATCTGAGAAGCAAAAGGACTTCTC 783
QY 255 HisSerAspGlnGlnLeu 260
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 784 CACTCGACCGACGACTG 801

```

RESULT 5

US-09-097-319A-5  
 ; Sequence 5, Application US/09097319A  
 ; Patent No. 6384207

GENERAL INFORMATION:  
 ; APPLICANT: Ainley, Michael  
 ; APPLICANT: Armstrong, Katherine  
 ; APPLICANT: Belmar, Scott  
 ; APPLICANT: Folkerts, Otto

```

APPLICANT: Hopkins, Nicole
APPLICANT: Menke, Michael A.
APPLICANT: Mayday, Dayakar
APPLICANT: Petolino, Joseph F.
APPLICANT: Smith, Kelley
APPLICANT: Woosley, Aaron
TITLE OF INVENTION: Regulatory Sequences for Transgenic Plants
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESS: Dowland Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/097,319A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317 337 4816
TELEFAX: 317 337 4847
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1354 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: Linear
MOLECULE TYPE: DNA
US-09-097-319A-5
Alignment Scores:
Pred. No.: 2.8e-76 Length: 1354
Score: 822.00 Matches: 168
Percent Similarity: 66.57% Conservative: 51
Best Local Similarity: 51.05% Mismatches: 90
Query Match: 51.25% Indels: 20
DB: 4 Gaps: 4
US-10-047-825-4 (1-313) x US-09-097-319A-5 (1-1354)
QY 5 ThrLeuMetGlnCysLeuValAlaValSerLeuSerCysValAlaHisAlaGln--- 23
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59 ACTTCATAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118
QY 24 -----LeuSerProThrPheTyrAlaSerSerCysProAsnLeu 36
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 CTGGCCACCATGGGCTGCTTCTTCCACAGTCTATGACCATTCCTCCAGGCGC 178
QY 37 GlnSerIleValAlaAlaMetThrGlnAlaValAlaSerGlnArgMetGlyAla 56
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 AAGAGATTGTCAGTGCATGTCAGTGCAGAGGCTGTGGCAAGAGACAGAGATGGCGCA 238
QY 57 SerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeu 76
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 239 TCTTTAGTCAGACGATTCCTTCATGACTGCTTGTCAAGGCTGCGCATGCTTGGGCTG 298
QY 77 LeuAspAlaGly-----GlyGluYstIrrAlaGlyProAsnLeuAsnSerVal 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 TTGACACAACGACGACCATATGATGAGAAAAGGCTCAACCGAAGCAAGATCTCCTC 358
QY 93 ArgGlyPheGluValIleAspThrIleYstArgAsnValGluAlaAlaCysProGlyVal 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 359 AGGGGTTTGAAGTGAAGCAAGATTAAGGCTGCTTGAAGCTGCTGCTGCCAGGAGACA 418
QY 113 ValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuGlyGly 132

```



```
Db 419 GTCCTCTGCGGACATGTTGCCCTTGGCTGCTGATTCACCGCCCTGTTGGTGA 478
|||
Qy 133 ProThrtPserValProLeuGlyArgArgAspSerThrtAlaSerAlaSerLeuAla 152
|||
Db 479 CCACTACTGGAGCGCTGCTGGCGGAGACGCTGCGGTGCAACATCCAGGGCTCC 538
|||
Qy 153 AsnSerAsnProProProThrtAlaSerLeuGlyThrtLeuIleSerLeuPheGlyArg 172
|||
Db 539 AACATGACATCCAGCCGCCAACACACACTCCCTCATCATCAGCAAGTTCAGCCG 598
|||
Qy 173 GlnGlyLeuSerProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyAla 192
|||
Db 599 CAGGGCTCAATGTTGATGTGTGCTGCGCTCTGAGTGTGATCAGCATTTGATGTCT 658
|||
Qy 193 ArgCysThrThrPheArgGlyArgIleTyr-----GlyAspThr 205
|||
Db 659 CGGTGCACTAGTTCCGGAGAGGCTATACAAACAGACAGGCAATGGCATGGCTGACAG 718
|||
Qy 206 AspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyGlyAsp 225
|||
Db 719 ACACGTGAGTATCCTACCGCCGCAAGGCTGAGGAGGATGCCCTGCTGTGTGAC 778
|||
Qy 226 GlyAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyrPheThr 245
|||
Db 779 AACACCTCTTCCCTTGACCTTCATCACCCCTGCCAAGTTTGACATTTTACTACAG 838
|||
Qy 246 AsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnGlnLeuPheAsnGlyGlySer 265
|||
Db 839 AACCTCTGCGCGGCAAGGCGCTTGAAGCTGTGATGTAGATTCGTACCAAGAGCGCT 898
|||
Qy 266 GlnAsp--AlaLeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsnAlaAspPhe 284
|||
Db 899 GACACGCGGCGCTGCTGAGGCAATGCTGCTGATGATCATCTCTTCCAGCACTTC 958
|||
Qy 285 ValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIle 304
|||
Db 959 GCACACTGCTATGTGTAATGTGGAAACATCTGCCCATGACAGGCTGACCAAGTGAGATC 1018
|||
Qy 305 ArgArgAsnCysArgValAlaAsnSer 313
|||
Db 1019 AGGAAGACTGCGAGGAGCTCAACAT 1045
|||
RESULT 6
US-09-615-192A-361
; Sequence 361, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Havukala, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615, 192A
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975, 316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713, 000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169, 789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 361
; LENGTH: 916
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-361
Alignment Scores:
Pred. No.: 5,65e-75 Length: 916
Score: 807.00 Matches: 162
Percent Similarity: 69.86% Conservative: 35
```

```
Best Local Similarity: 57.45% Mismatches: 75
Query Match: 50.31% Indels: 10
Db: 4 Gaps: 3
US-10-047-825-4 (1-313) x US-09-615-192A-361 (1-916)
Qy 9 CysLeuValAlaValSerLeuLeuSerCys-----ValAlaHisAlaGlnLeu 24
|||
Db 70 TGCATTCCTGTCATTTGTTGTTCAATATCTGTCATTAAGTATGATGATGATGAGGAACTG 129
|||
Qy 25 SerProThrPheThrAlaSerSerCysProAsnLeuGlnSerIleValArgAlaIleMet 44
|||
Db 130 AGCTCCAGCTTTATGATGAATGTTTCCCAAGCCCTGTCTATGATGCAAGCGGGGTG 189
|||
Qy 45 ThrGlnAlaValAlaSerGlnArgMetGlyAlaSerLeuLeuArgLeuPheHis 64
|||
Db 190 AAGCAACAGTGGCTGAAGAAACATATGAGGAGGATGCTTCCGCTCATTTCCAC 249
|||
Qy 65 AspCysPheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGly----- 80
|||
Db 250 GACTGCTTGTGTAATGAGCTGCGATGGCTGTGACTGTGGAATTCACAGACCTTCACT 309
|||
Qy 81 GlyIleuPsrThrAlaGlyProAsnLeuAsnSerValArgGlyPheGluValIleAspThr 100
|||
Db 310 AGCAGAAATATATGCTCTTCCCAATATACATTCGCGAGGGGTTTGAGAGTATGATAGC 369
|||
Qy 101 IleuArgAsnValGluAlaAlaCysProGlyValAlaSerCysAlaAspIleLeuAla 120
|||
Db 370 ATTAAGAGCCACTCGAAGTGTGCTCACCGCGCTGCTTCTTGACACATTCCTCAG 429
|||
Qy 121 LeuAlaIlaArgAspGlyThrAsnLeuLeuGlyGlyProThrtPserValProLeuGly 140
|||
Db 430 ATGCTCTGCTGATCTGTTCTGTTGATGTTGGTGACCTTGTGAGAGTATGATGAGG 489
|||
Qy 141 ArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThr 160
|||
Db 490 AGGCAAGACTCAACACAGGAGCATTTAGCGGTGCACAAATACATTCGCCCTCCACT 549
|||
Qy 161 AlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMet 180
|||
Db 550 TCCATCTGACGAAACTCATTTATTTTCAAGGCCACAGGCGCTCCACAAAGGAATG 609
|||
Qy 181 ThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArgGlyArg 200
|||
Db 610 GTTCCACTCTGCTGCTGCTATACATGCGGAGGCGCAATTCAGAGATTCAGAGCCAT 669
|||
Qy 201 IleTyrGlyAspThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnIleThrCysPro 220
|||
Db 670 ATTACAAACGACACACATAGATGATGATGATGATGATGATGATGATGATGATGATG 729
|||
Qy 221 ArgSer-----GlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrProValArg 238
|||
Db 730 AGTACACAGGCTCGGAGACACACCTGCGCCCTGATTAAGACATCCACTGTG 789
|||
Qy 229 PheAspThrAlaTyrPheThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAspGln 258
|||
Db 790 TTGACAAAACATTTTCTATCATCTGAAAGCAAAAGAGACTTCACACCGACAG 849
|||
Qy 259 GluLeuPheAsnGlyGlySerGlnAspAlaLeuValArgGlnTyrSerAlaSerAlaSer 278
|||
Db 850 GAACCTTCAACAGGAGCTCAGTATGATGATGATGATGATGATGATGATGATGATGAT 909
|||
Qy 279 LeuPhe 280
|||
Db 910 ACCTTC 915
|||
RESULT 7
US-09-365-150-3
; Sequence 3, Application US/09365150
; Patent No. 6278041
; GENERAL INFORMATION:
; APPLICANT: Lagrimini, Mark
; APPLICANT: Desai, Nalini
```



Oy 30 AlaSerCysProAsnIleuSerIleValAlaAlaAlaMetThrGlnIleValAla 49  
 Db 102 GATACACATTCGCCCTTAATGTTACAGATTAATTTGTACGTGCTGTTATGTCATCAAGCAACGT 166  
 Oy 50 SerGluGlnArgMetCylAlaSerIleuLeuArgLeuPheHisAspCysPheValGln 69  
 Db 162 ACTGATGCTCGAGCTGGTGCTAAATATTCGTCTTCATTTCCATGATTTGTTGTAAAT 221  
 Oy 70 GlyCysArgIleSerIleLeuLeuAspAlaGlyIly-----GluysThrIleIly 86  
 Db 222 GGTGTGAGTACATCAATTTGTTGAGACACAGATGGGCTCAACTGAGAAATACACT 281  
 Oy 87 ProAsnLeuAsnSerValArgIlyPheGluValIleAspThrIleLysArgAsnIleGlu 106  
 Db 282 GCTAATGTAGTGCA---GAGAGATTTGATTTGTGATGATATTAACTGCCTAGAG 333  
 Oy 107 AlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaArgAspIly 126  
 Db 339 AAGTATGCCCTGGTGTTGATCTTGCCACATATTTTATCCCTTCATCAATAATTGA 398  
 Oy 127 ThrAsnLeuGluGlyIlyProThrIlePheValProLeuGlyIlyArgArgAspSerThr 146  
 Db 399 GTTGTCTGGCTAAGAGTCCTGCGCGCAAGTACTTTTGGCAGAAAGATAGCTTACA 455  
 Oy 147 AlaSerAlaSerLeuAlaAsnSerAsnProProProThrAlaSerLeuGlyThrIleu 166  
 Db 459 GCAACCGATCTGCAGCTAATGTATTCCTGACCCCTTGAAACCCCTTCCTGATAG 518  
 Oy 167 IleSerIlePheIlyArgIleIlyLeuSerProArgAspMetThrIleLeuSerGlyAla 186  
 Db 519 ACGCCCAATTCACCAACAGAGGAATTTAACTACTCTGTCCTATCTGTGGTGA 578  
 Oy 187 HisThrIleGlyGlnIleArgCysThrThrPheArgIlyArgIleIly----- 207  
 Db 579 CACACATTTGGAAGACCAAGATGTGTACTTTTGACACAGCTCTTTAACTCACTAGTGC 633  
 Oy 204 -----AspThrAspIleAsnIleAspPheAlaIleuArgGlnGlnThrCys 211  
 Db 639 AGTGGTAATCCTATCCAAACGCTAAGCCTCATTTTATCAAACTTACGGGGAATTTGT 698  
 Oy 220 ProArgSerGlyIlyAspIlyAsn---LeuAlaProIleAspValGlnThrProValArg 239  
 Db 699 CCTCAAGTGGAAATATATGGCAATACCTTTCAAAATCTGTATATATCAACTCCTATATGT 758  
 Oy 239 PheAspThrAlaTyrrPheThrAsnLeuLeuSerArgArgIlyLeuPheHisSerAspGln 258  
 Db 759 TTTGATATGACATTTTTCACATCTTCAATCTTCAAAATATCAAGAGACTTCTTCAAACTGATCA 817  
 Oy 259 GluLeuPhe-----AsnGlyIlySerGlnAspAlaLeuValArgGlnIlyrSerAlaSer 278  
 Db 819 GAGTGTGTTTCTCATCTGTGATCTGTACAAATTCGTATATGGAATCGTATATGCTGTAGT 877  
 Oy 277 AlaSerLeuPheAsnIleAspPheValAlaAlaMetIleArgMetGlyAsnValGlyAl 296  
 Db 879 CAAGTCAGTTTGTGATGATTTTATTGTCGAGATTAATATGATATTAATATCA 933  
 Oy 297 LeuThrGlyThrAlaGlyGlnIleArgArgAsnCysArgValAlaAsn 312  
 Db 939 TTAACGTGCTAATGAGAGATTTAGGAAGATTTCAAGACAGGTTAAT 986

STREET: P.O. Box 2113  
City: Madison .  
STATE: WI  
COUNTRY: USA  
ZIP: 53701-2113  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Vers1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,094  
FILING DATE: 03-MAR-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/378,588  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27,386  
REFERENCE/DOCKET NUMBER: 11-229-9101-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 251-2484  
TELEFAX: (608) 251-9166  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1270 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-811-094-23

Alignment Scores:	
Pred. No.:	7,45e-66
Score:	721.50
Percent Similarity:	63.978
Best Local Similarity:	47.768
Query Match:	44.988
DB:	2
Length:	1270
Matches:	151
Conservative:	51
Mismatches:	97
Indels:	17
Gaps:	6

```

US-10-047-825-4 (1-313) x US-08-811-094-23 (1-1270)

Oy      10  LeuValAlaIalSerLeuLeuSerCysValAlaHisAlaGlnLeuSerProThrPheTy 29
          |||||
          51  TTTGGTCGAATTTTGGCAGCATC-----AATCTTAATTAACGCCACATTTTAC 101
          |||||

Oy      30  AlaSerCysProAsnLeuGlnSerIleValArGlaIalaMetThrGlnAlaVala 49
          :|||
          102  GATACCACTTCCCTCAATAGTTTACAAGATTTGTACGTGGTGTATGGATCAAAAGCAACGT 161
          |||||

Oy      50  SerGluGlnArGmetGlyAlaSerLeuLeuArGleuPhePheHisApyCysPheValGln 69
          |||||
          162  ACTGATGCTGCAGCGCGTGCTAAATTAATTCGCTTCATTTCCATGATGTTTGTGTAAT 221
          |||||

Oy      70  GlyCysAspGlySerIleLeuLeuAspAlaGlyIly-----GluTyThrAlaGly 86
          |||||
          222  GGTGTGTGATGCATCAATTTGTTAGACACAGATGGAGCTCAAACTGAGAAGAAGTGCACCT 281
          |||||

Oy      87  ProAsnLeuAsnSerValArGlyPheGluValIleAspThrIleTyAsrAsnValGlu 106
          |||||
          282  GCTAATGTAGAGTGCC--GGAAGGATTTGATATGTGGATGATTTTAACTGCACCTAGAC 338
          |||||

Oy      107  AlaAlaCysProGlyValAlaIalSerCysAlaAspIleLeuAlaLeuAlaIalArGaspGly 126
          |||||
          339  AATGATATGCCCGGTGGTGTATCTTGTGCACAAATTTTATAGCCCTGCATCTGAAATTTGGA 398
          |||||

Oy      127  ThrAsnLeuLeuGlyIlyProThrTrpSerValProLeuGlyArGArGAspSerThrThr 146
          |||||
          399  GTTGTCTGGCGCTAAAGGTCCGTCGGCGCAACACTTTTGGCCAAACAAGATAGCTTACA 458
          |||||

Oy      147  AlaSerLeuSerLeuAlaAsnSerAsnProProProThrAlaSerLeuGlyThrLeu 166
          |||||

```

Db 459 GCAACCGATCTGAGCTAATAGTATCTTACGCCCCCTTGAAACCTTGCTGTAAAG 518  
 QY 167 IleserLeuphediArgngInglyLeuSerProArgAspMethrAlaLeuSerGlyAla 186  
 Db 519 ACCGACAAATTCACCAACGAGGAAATGATTTACTGATCTTCTCTATCTGCTGCA 578  
 QY 187 HisThrIleGlyAlaIleArgCysThrThrPheArgGlyArgIleGly----- 203  
 Db 579 CACACATTTGGAGAGCAGAGATGTGCTTGAACAACGCTCTTAACCTCAAGTGC 638  
 QY 204 -----AspThrAspIleAsnAlaSerPheAlaIleLeuArgInglyInthrCys 219  
 Db 639 AGTGGAATTCCTGATCCACCGAGAGCTACTTTTTCACAAACATTCAGCAATTTGT 698  
 QY 220 ProArgSerGlyGlyAspGlyAsn---LeuAlaProIleAspValGlnThrProValArg 238  
 Db 699 CCTCAAGGTGGAATATATGCAATCTTTTACAAATCTTGATATATCAACTCTCTAATGAT 758  
 QY 239 PheAspThrAlaIleArgPheThrAsnLeuLeuSerArgArgGlyLeupheHisSerAspGln 258  
 Db 759 TTTGATTAATGACTATTCTCACTATCTTCAAAATATATCAAGAGACTTCTCAACATGATCA 818  
 QY 259 GluLeuphe-----AsnGlyGlySerGlnAspAlaLeuValArgGlnThrSerAlaSer 276  
 Db 819 GACTTTCTTTCTACATCTGAGATCTGCTACAAATTCATATGTAATGCTATCTGCTGAT 878  
 QY 277 AlaSerLeupheAsnAlaAspPheValAlaIleMetIleArgMetGlyAsnValGlyVal 296  
 Db 879 CAAAGTCAGTTTCTTATGATATTTTATTTGCTGCGATGATTAATGCTAATATGATCA 938  
 QY 297 LeuThrGlyThrAlaGlyInglyInleArgArgAsnGlyArgValValAsn 312  
 Db 939 TTTACTGTTACTTAATGAGATTAAGAAATGATTCAGAAAGGTTAAT 986  
 RESULT 10  
 PCT-US94-11121-23  
 : Sequence 23, Application PC/TUS9411121  
 : GENERAL INFORMATION:  
 : APPLICANT: John, Maliyakal E.  
 : TITLE OF INVENTION: TRANSGENIC COTTON PLANTS  
 : TITLE OF INVENTION: PRODUCING HETEROLOGOUS PEROXIDASE  
 : NUMBER OF SEQUENCES: 23  
 : CORRESPONDENCE ADDRESSES:  
 : ADDRESSEE: Nicholas J. Seay, Charles & Brady  
 : STREET: First Wisconsin Plaza, One South  
 : STREET: Plackney St.,  
 : CITY: Madison  
 : STATE: WI  
 : COUNTRY: USA  
 : ZIP: 53701-2113  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US94/11121  
 : FILING DATE:  
 : CLASSIFICATION:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Seay, Nicholas J.  
 : REGISTRATION NUMBER: 27,386  
 : REFERENCE/DOCKET NUMBER: 11-229-9076-8  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (608) 251-2484  
 : TELEFAX: (608) 251-9166  
 : INFORMATION FOR SEQ ID NO: 23:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1270 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear

: MOLECULE TYPE: DNA (genomic)  
 : PCT-US94-11121-23  
 Alignment Scores:  
 Pred. No.: 7,45e-66 Length: 1270  
 Score: 721.50 Matches: 151  
 Percent Similarity: 63.92% Conservative: 51  
 Best Local Similarity: 47.78% Mismatches: 97  
 Query Match: 44.98% Indels: 17  
 DB: 5 Gaps: 6  
 US-10-047-825-4 (1-313) x PCT-US94-11121-23 (1-1270)  
 QY 10 LeuValAlaValSerLeuLeuSerCysValAlaHisAspGlnLeuSerProThrPheThr 29  
 Db 51 TTGGTGAATTTTGGCAGCATCA-----AATGCTCAATTAACGCCAATTTTAC 101  
 QY 30 AlaSerSerCysProAsnLeuGlnSerIleValArgAlaIleMetThrGlnAlaValAla 49  
 Db 102 GATACCACTTCCCTTAATGTTACAAAGTATTTAGCTGTTGTTATGATCAAGCAACGT 161  
 QY 50 SerGluGlnArgMetGlyAlaSerLeuLeuArgLeuphePheHisAspCysPheValGln 69  
 Db 162 ACTGATGCTGAGCTGGTGGCTAAATATTCGTTCTCATTTCCATGATTTGTTGTAAT 221  
 QY 70 GlyCysAspGlySerIleLeuLeuAspAlaGlyGly-----GluCysThrAlaGly 86  
 Db 222 GGTGTGATGATGATCAATTTTCTTAGACACAGATGGGACTCAACTGMAAGATGCACT 281  
 QY 87 ProAsnLeuAsnSerValArgGlyPheGluValIleAspThrIleCysArgAsnValGlu 106  
 Db 282 GCTATGTAGTGCAGCA---GGAGGATTTGATATGTCATATTAACCTGCACATGAG 338  
 QY 107 AlaIleCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaIleArgAspGly 126  
 Db 339 AATGATATGCCCTGGTGTGTTGTTGTCAGATTTTACCTTGATCTGAAATTTGA 398  
 QY 127 ThrAsnLeuGlyGlyProThrTrpSerValProLeuGlyArgArgAspSerThrThr 146  
 Db 399 GTTCTTGGCTTAAGGTCGTCGTCGTCGCAAGTACTTTTGGCAGAGAGATGACTTAACA 458  
 QY 147 AlaSerAlaSerLeuAlaAsnSerAsnProProProThrAlaSerLeuGlyThrLeu 166  
 Db 459 GCAACCGATCTGAGCTAATAGTATTCCTGAGCCCTTTGAACCCCTTGCTGTAATG 518  
 QY 167 IleserLeuphediArgngInglyLeuSerProArgAspMethrAlaLeuSerGlyAla 186  
 Db 519 ACCGACAAATTCACCAACGAGGAAATGATTTACTGATCTTCTCTATCTGCTGCA 578  
 QY 187 HisThrIleGlyAlaIleArgCysThrThrPheArgGlyArgIleGly----- 203  
 Db 579 CACACATTTGGAGAGCAGAGATGTGCTTGAACAACGCTCTTAACCTCAAGTGC 638  
 QY 204 -----AspThrAspIleAsnAlaSerPheAlaIleLeuArgInglyInthrCys 219  
 Db 639 AGTGGAATTCCTGATCCACCGAGAGCTACTTTTTCACAAACATTCAGCAATTTGT 698  
 QY 220 ProArgSerGlyGlyAspGlyAsn---LeuAlaProIleAspValGlnThrProValArg 238  
 Db 699 CCTCAAGGTGGAATATATGCAATCTTTTACAAATCTTGATATATCAACTCTCTAATGAT 758  
 QY 239 PheAspThrAlaIleArgPheThrAsnLeuLeuSerArgArgGlyLeupheHisSerAspGln 258  
 Db 759 TTTGATTAATGACTATTCTCACTATCTTCAAAATATATCAAGAGACTTCTCAACATGATCA 818  
 QY 259 GluLeuphe-----AsnGlyGlySerGlnAspAlaLeuValArgGlnThrSerAlaSer 276  
 Db 819 GACTTTCTTTCTACATCTGAGATCTGCTACAAATTCATATGTAATGCTATCTGCTGAT 878  
 QY 277 AlaSerLeupheAsnAlaAspPheValAlaIleMetIleArgMetGlyAsnValGlyVal 296  
 Db 879 CAAAGTCAGTTTCTTATGATATTTTATTTGCTGCGATGATTAATGCTAATATGATCA 938

Oy	297	LeuThrlGlyThrAlaGlyGlnIleArgAspHisValValAsn	312
Dd	939	TTrAcTGtCTrACTrATrAGcAGcATTTrAGAAgATTGCACAgGGTTAAr	986
RESULT 11			
US-09-365-150-1			
Sequence 1, Application US/09365150			
Patent No. 6278041			
GENERAL INFORMATION:			
APPLICANT: Lagrimini, Mark			
APPLICANT: Desai, Nalini			
TITLE OF INVENTION: No. 6278041el Peroxidase Gene Sequences			
FILE REFERENCE: S-31081PI			
CURRENT APPLICATION NUMBER: US/09/365,150			
CURRENT FILING DATE: 1999-07-30			
NUMBER OF SEQ ID NOS: 6			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 1			
LENGTH: 975			
TYPE: DNA			
ORGANISM: Nicotiana glaucaformis			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (1)..(972)			
US-09-365-150-1			
Alignment Scores:			
Pred. No.:	2,67e-65	Length:	975
Score:	714.50	Matches:	149
Percent Similarity:	63.92%	Conservative:	53
Best Local Similarity:	47.15%	Mismatches:	17
Query Match:	44.54%	Indels:	97
DB:	4	Gaps:	6
US-10-047-825-4 (1-313) x US-09-365-150-1 (1-975)			
Oy	10	LeuValAlaValSerLeuLeuSerCysValAlaHisAlaGlnLeuSerProThrPhetyr	29
Dd	37	TTGGTTCACAAATTTTTCGACGATCA-----AATGCTCAATTAGAATCCCATTTTAC	87
Oy	30	AlaSerSerCysProAsnLeuGlnSerIleValArgAlaAlaMetThrGlnAlaValala	49
Dd	88	GATGACCTGTTCCTTAAGTTACAAGATATGTACGTGGTAGTTATGATCAAGCAAGCGT	147
Oy	50	SerGlnGlnArgMetGlyAlaSerLeuLeuArgLeuPheHisAspCysPheValGln	69
Dd	148	ACGTATGCTCGACGTGGTGTCAAATTTATTCGTCCTTCATTTCCAGATTCCTTGCAT	207
Oy	70	GlyCysAspGlySerIleLeuLeuAspAlaGly-----GlnLysThrAlaGly	86
Dd	208	GGTTGTGATGGATGCCATTTTATTAACACAGATGGGACTCCAACATGAGAAAGATCACCT	267
Oy	87	ProAsnLeuAsnSerValArgIlyPheGluValIleAspThrIleLysArgAsnValGlu	106
Dd	268	CCTTAATGTAGCTSC---GGAAGATTTGATATTGTTGATGATATTTAAACTGCATTAAG	324
Oy	107	AlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspLy	126
Dd	325	AATGTATGCCCTGTGGTGTGTATCTTGCGAGATAATTTTATCCCTTGCAATCAATTAAG	384
Oy	127	ThrAsnLeuLeuGlyGlyProThrTrpSerValProLeuGlyValArgArgAspSerThr	146
Dd	385	GTTCGCTTGGCGAAGGTCCTGCATGCGACATCTTTTGGCAGAAAACAGCTTAACA	444
Oy	147	AlaSerAlaSerLeuAlaAsnSerAsnProProProProThrAlaSerLeuGlyThrLeu	166
Dd	445	GCAAAACGATCTGAGCGTAATAAGATATCCCCAGACCCTTTGAAACCCCTGCTATATG	504
Oy	167	IleSerLeuPheGlyValGlnGlyLeuSerProArgAspMetThrAlaSerSerVala	186
Dd	505	ACACCACTATTCACCAACAGAGGAATGATTTAACATGATCTTGTTCATACAGGTSCA	564
Oy	187	HisThrIleGlyGlnAlaArgCysThrThrPheArgGlyValGllLeTyely-----	203

Db	565	CATGACATTTGGAGACGACGAGATGTGTGACTTTTGACACAGCTCTCTTAACTTCAGTGGC	62
Oy	204	-----AspThrAspIleAsnLaserPheAlaLeuArgIleGlnIthrCys	211
Db	625	AGTGGTACCCTGATCCCTACCCCTACAGCCTCAATTTTACAAACATTACAGGAATTTGT	68
Oy	220	ProArgSerGlyGlyValSpgIysn---LeuAlaProIleAspValGlnProValArg	231
Db	665	CCTACAGCTGGAAATATATGGCAATCTTTACAAATCTGTATATATCACTCTATATGC	74
Oy	239	PheAspThrAlaTyrPheThrAsnLeuLeuSerArgArgIlyLeuPheIleSerAspGln	25
Db	745	TTTGTATATGACTATTTTCCAAACCTTCMAATATATCAGGGCTCTTCAACVGTATCA	80
Oy	259	GluLeuPhe-----AsnGlyGlySerGlnAspAlaLeuValArgGlnTyrSerAlaSer	27
Db	805	GAGTGTGTTTCAACATCTGATCTCTCAATTCGATAGTATATCGTTATGTGTGTAC	86
Oy	277	AlaSerLeuPheAsnAlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyVal	29
Db	865	CAAACTGAGTTTGTGATGATTTTGTAGCTCTATGATTAATTTGGTATATATAGTCA	92
Oy	297	LeuArgIlyThrIlaGlyGlnIleLeuArgArgAsnSarGlyValAlaAsn	312
Db	925	TTAACTGTACTAATGAGAAATTAGACAGATTTCAGAGGGTTAAT	972
	RESULT 12		
	US-09-365-150-4		
	Sequence 4, Application US/09365150		
	Patent No. 6278041		
	GENERAL INFORMATION:		
	APPLICANT: Lagrimini, Mark		
	APPLICANT: Desai, Nalini		
	TITLE OF INVENTION: NO. 6278041el Peroxidase Gene Sequences		
	FILE REFERENCE: S-31081P1		
	CURRENT APPLICATION NUMBER: US/09/365,150		
	CURRENT FILING DATE: 1999-07-30		
	NUMBER OF SEQ ID NOS: 6		
	SOFTWARE: PatentIn Ver. 2.0		
	SEQ ID NO 4		
	LENGTH: 975		
	TYPE: DNA		
	ORGANISM: Artificial Sequence		
	FEATURE:		
	OTHER INFORMATION: Description of Artificial Sequence: Synthetic N.		
	OTHER INFORMATION: CommentIsIformis peroxidase gene		
	US-09-365-150-4		
	Alignment Scores:		
	Pred. No.:	2, 67e-65	975
	Score:	714.50	149
	Percent Similarity:	63.92%	Conservative: 53
	Basic Local Similarity:	47.15%	Mismatches: 97
	Query Match:	44.54%	Indels: 17
	Db:	4	Gaps: 6
	US-10-047-825-4 (1-313) x US-09-365-150-4 (1-975)		
Oy	10	LeuValAlaValSerLeuLeuSerCysValAlaIleAsnIleuSerProThrPheTyr	29
Db	37	CTCTGTTCGCAATCTTCGCCGCCGACG-----AACGCCAGCTGAGCGCACACTTCTAC	87
Oy	30	AlaSerSerCysProAsnLeuGlnSerIleValArgAlaAlaMetThrGlnAlaValAla	49
Db	88	GAGCAGCACTGCCCCCAAGCTGACGACGCACTTCGCGCGCGGTAGAGGACACGCCCGC	14
Oy	50	SerGluGlnArgMetGlyAlaSerLeuLeuArgLeuPhePheIleAspCysPheValGln	69
Db	148	ACGAGCGCGCGCGCGGCGGACCAAGATCATCCGCTGCACCTCCACGACGACTCTTCGTAAC	20
Oy	70	GlyCysAspGlySerIleLeuLeuAspAlaGly-----GluysThrAlaGly	86

```

Db 208 GGCCTGCACGCGACATCTCTGACACGACGCGACACCGACGAGACGCTGCT 267
Qy 87 ProasnuLeuanserValarglyPhegluValIleasPthrllelyArgAsnValglu 106
Db 268 CCGAACGCTGGGCGCC--GGCGGCTTGCACATCTGTGGACATGACATGAACCGCCCTCGAG 324
Qy 107 AlaIacysProgllyValIasercysAlaasPleleuAlaIleuAlaIaArgaspIly 126
Db 325 AACGTGTCCCGGAGTGTGAGCTGCGCGACATCTGTGAGCTGCGAGAGATGCGC 384
Qy 127 ThrAsnuLeuenglyProthrPpserValProleuGlyArgArgAspserThrThr 146
Db 385 GTGGCGCTGGCCGACGCGCTTGTGACAGAGCTGCTTGTGCGAGAGACGCTGACT 444
Qy 147 AlaserAlaserLeuAlaasnsersAsnProProProProThrAlaserLeuGlyThrleu 166
Db 445 GCCAACGCGACGACGACGACGACGACATCCCGACGCGCTTGCAGACCGCCGCGTGTG 504
Qy 167 IlaserLeupheglYargInglYleuserProArgAspMetThrAlaIleuSerGlyAla 186
Db 505 ACCCCCTGTTCCACCAACAAAGGCGATGACCTGACCGACCTGCTGGCCAGAGCGGCGC 564
Qy 187 HisThrIleaglAlaIaargCysThrThrPheArgglyArgIleTyrgly----- 203
Db 565 CACACCTTGGGTGCGAGCCGCTGCGGACCTTGCAGACAGACAGACTTCACTCAGCGGC 624
Qy 204 -----AspThrAspIleasnaIasersPheAlaIaIleuArgInglThrCys 219
Db 625 AGCGGCAACCCCGACCCGACCTGTGACGCCACCTTCTTCCAGACCTGAGGCGATCTGC 684
Qy 220 ProArgSerIleGlyAspIlyAsn---LeuAlaProIleAspValglThrProValArg 238
Db 685 CCCGAGCGCGGCAACAGCGACACCTTCCACCAACCTGACATGACGACGACCCGACGAC 744
Qy 239 PheAspThrAlaIyrPheThrAsnuLeuLeuserArgArgglyLeuPheHisSerAspGln 258
Db 745 TTCGACAAACGACTTCTTGCACCACTCTCGAAGAACCAAGGCGCTCTCCAGACGACG 804
Qy 259 GluleuPhe-----AsnglyIySerGlnAspAlaIleuValArgInglTySerAlaser 276
Db 805 GAGCTGTTCGACACACGCGGACGCGCCACCATCGCCATGTGACCTGACGCGCGACG 864
Qy 277 AlaserLeupheAsnaIasPheValAlaIaIleuMetIleArgMetGlyAsnValglYal 296
Db 865 CAGACCCAGTCTTCTGACGACCTTGTGACGACATGATGAACCTGGCAACATCTGCTCC 924
Qy 297 LeuThrIleGlyAlaIaIleuGlnIleArgArgAsnCysArgValYalAsn 312
Db 925 CTGACCGGCGACCAAGCGGAGATCGGACGACGACTGCAACGCGGTGAC 972

RESULT 13
US-09-615-192A-371
: Sequence 371, Application US/09615192A
: Patent No. 6410718
: GENERAL INFORMATION:
: APPLICANT: Bloksberg, Leonard N.
: APPLICANT: Havukkala, Ilkka
: TITLE OF INVENTION: Materials and Methods for the
: FILE REFERENCE: 11000.10034U
: CURRENT APPLICATION NUMBER: US/09/615,192A
: CURRENT FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 08/975,316
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: US 08/713,000
: PRIOR FILING DATE: 1996-09-11
: PRIOR APPLICATION NUMBER: US 09/169,769
: PRIOR FILING DATE: 1998-10-09
: NUMBER OF SEQ ID NOS: 405
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 371
: LENGTH: 1522
: TYPE: DNA

```

```

: ORGANISM: Pinus radiata
US-09-615-192A-371

Alignment Scores:
Pred. No.: 9.55e-60 Length: 1522
Score: 664.00 Matches: 146
Percent Similarity: 64.97% Conservative: 58
Best Local Similarity: 46.50% Mismatches: 98
Query Match: 41.40% Indels: 13
DB: 4 Gaps: 5

US-10-047-825-4 (1-313) x US-09-615-192A-371 (1-1522)

Qy 10 LeuValAlaValSerLeuLeuSerCysValAlaIleuSerProthrPheThr 29
Db 141 ATGGCGCGCTGTGAATGCTTGCCTGCGGCTG- GCTGTCGTCTTGTGACGCTTCTAC 199
Qy 30 AlaserSerCysProasnuLeuGlnSerIleValArgAlaIleuMetThrGlnAla 49
Db 200 AGCTGAGTTCCTCCGCTTGTGAGTCCATGATGTGGAGCGCATGAGACCTATTGAGT 259
Qy 50 SerGlnIaArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAspCysPheValGln 69
Db 260 GCAGACATCACACAGCGCTGCGAGATGTTGTGAGGCTCCACTCCAGACGCTTGTCCAG 319
Qy 70 GlyCysAspGlySerIleLeuLeuAspAla-----GlyGlyGlnIyThrAlaGlyProasn 88
Db 320 GGTATCGATGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 379
Qy 89 LeuAsn---SerValArgglyPhegluValIleasPthrllelyArgAsnValgluAla 107
Db 380 TTATCATCAGACGCGAGCGCTTAAAGATATTAAACGACATGAAGAGACGCTGACAGCC 439
Qy 108 AlacysProgllyValIasercysAlaasPleleuAlaIleuAlaIaArgaspIlyThr 127
Db 440 GCTTACAGCGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 499
Qy 128 AsnuLeuenglyIyProthrPpserValProleuGlyArgArgAspserThrThr--- 146
Db 500 GTAATGCTGGAGGACGCTTCTTACCCCTTACCTGCGCGGACGACACCTTACCTTC 559
Qy 147 AlaserAlaserLeuAlaasnsersAsnProProProProThrAlaserLeuGlyThrleu 166
Db 560 GCGAATGATGACCGCTTCTGCGCAATTTGCCATTCGCCAATCTCCCAATGACGCGCTC 619
Qy 167 IlaserLeupheglYargInglYleuserProArgAspMetThrAlaIleuSerGlyAla 186
Db 620 ATCAGCTGTTGGGTGCCCAAGGCTTGAATTTGCAGAGATCTGTGCGCCCTCTCAGAGA 679
Qy 187 HisThrIleaglAlaIaargCysThrThrPheArgglyArgIleTyrgly----- 203
Db 680 CATACAAATGGCACACCACTGCTCTCTTGTGACACAGCATATTAACGACGACCC 739
Qy 204 -----AspThrAspIleasnaIasersPheAlaIaIleuArgInglThr 218
Db 740 GGTACACAAATGCGGATCCCGACGATGACGACGACGATTTGCTCAAAAGCTTATCTACCC 799
Qy 219 CysProArgSerGlyIyAspGlyAsnLeuAlaProIleAspValglThrProValArg 238
Db 800 TCCCTTACCACT---ACCACCGTTTAAACACCAACCAATTTGATATTCGACTCCAAATGTG 856
Qy 239 PheAspThrAlaIyrPheThrAsnuLeuLeuserArgArgglyLeuPheHisSerAspGln 258
Db 857 TTCGACAAACAAATACATGCTGATCTCTCCACCAAGCGACGACGCTCTTCACTTACAG 916
Qy 259 GluleuPheAsnglyIySerGlnAspAlaIleuValArgInglTySerAlaserThr 278
Db 917 ACTCTTACACCGACACTCGAACCCCGGACATTTGATGAATTTGCGGTGATATGAGC 976
Qy 279 LeuPheAsnaIasPheValAlaIaIleuMetIleArgMetGlyAsnValglYalIleuThr 298
Db 977 CTCTTCTTGAACAGTGTGTGTGACGACATGCTCAAAATGCGGACGCTGAGTGTCTACA 1036

```





```
QY      82  GluylsThrAlaGlyProAsnLeuAsnSerValAlaArglyPheGluValIleAspThrIle 101
      ||||| : : : ||||| ||||| : : : ||||| ||||| ||||| ||||| ||||| |||||
Db      5107 GAGAAAGGCTCAACCCAGACAGAACTCTCCAGGGGGTTGAGTGCACACAGATT 5166
QY      102  LysArgAsnValAlaIleAlaIleAsnProGlyValIleSerCysAlaAspIleLeuAlaLeu 121
      ||| : : : ||||| ||||| : : : ||||| ||||| : : : ||||| ||||| : : : |||||
Db      5167 AAGGCTCTCTGAGGCTGCTGCCAGCAGCAGCTCTCTGCGGACATGTTGGCCCT 5226
QY      122  AlaIleArgAspIleThrAsnLeu----- 129
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5227 GCGGCTGTGATTCACGCCGCCGTGATGTCACACTATGCACAACTCTTCCAACTCA 5286
QY      129 ----- 129
Db      5287 GGAACAGACATGATATTGTGTGTGTGTGTATATAATATATAGTAGACTT 5346
QY      130 ----- 130
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5347 GGCAAACTGATATTCTTCTGAGCTCTAAACCGTGGTGGGACCTACTGAGGACGTG 5406
QY      138  ProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProPro 157
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5407 CCACTGGCCGAGAGACTGCTCGGTGCACATCCAGGGCTCCAAACATGATGCCA 5466
QY      158  ProProThrAlaSerLeuGlyThrIleLeuSerLeuPheGlyArgGlyGlyLeuSerPro 177
      ||| : : : ||||| ||||| : : : ||||| ||||| : : : ||||| ||||| : : : |||||
Db      5467 GCCCCCAACAACACACTCCCGCTATCATCCAGATTCACAGCCGAGGCCCTCAATGTT 5526
QY      178  ArgAspMetThrAlaLeuSerGlyAla----- 186
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5527 GTTGATGTGTGCGCCCTCAGGTGA-TTTTCTGTATTATTATGATCATCTGTCTT 5585
QY      186 ----- 186
Db      5586 CGTTATTCACCAACTTACGCGACACTCATATTACCATGATACATATCATCTGTCAAT 5645
QY      187 ----- 187
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5646 ACAGGTGCTACACACATGATGTCTCGTGCCTAGTTCGCGCAGAGGCTATACAC 5705
QY      203 ----- 203
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5706 CAGACAGGCAATGCGATGCTGCACACACTGATGATCTTACGCCGCAAACTGAGG 5765
QY      217  GluThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGluThrPro 236
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5766 CAGGAGATGCCCGCTGTGTGTGACAAACACTCTTCCCTGACTTCATCAGCCCT 5825
QY      237  ValArgPheAspThrAlaArgPheThrAsnLeuLeuSerArgGlyLeuPheHisSer 256
      : : : ||||| ||||| : : : ||||| ||||| : : : ||||| ||||| : : : |||||
Db      5826 GCCAAGTTTGACATTTTACTACAGAACCTCTGCGCGGAGAGGCCCTCTAGGCT 5885
QY      257  AspGluGluLeuPheAsnGlyGlySerGlnAsp---AlaLeuValArgGlyLeuSer 275
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5886 GATGAGATTTCTGTTAACCAAGAGCGCTGAGACAGCGCCCTGTAAGGCTATGAGCT 5945
QY      276  SerAlaSerLeuPheAsnAlaAspPheValAlaIleMetIleArgMetGlyAsnValGly 295
      : : : ||||| ||||| : : : ||||| ||||| : : : ||||| ||||| : : : |||||
Db      5946 GATGTCATCTCTTCTCCAGCAGCTTGCACAGCTGATAGTGGAATATGGAACAATCTCG 6005
QY      296  ValLeuThrGlyThrAlaGlyGluIleArgArgAsnCysArgValValAsnSer 313
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      6006 CCACTGACAGGCTGATGAGTGAAGAACTCAGAGAGCTCAACAAAT 6059

RESULT 15
US-09-615-192A-365
: Sequence 365, Application US/09615192A
: Patent No. 6410718
: GENERAL INFORMATION:
: APPLICANT: Bloksberg, Leonard N.
: APPLICANT: Havukkala, Ilkka
: TITLE OF INVENTION: Materials and Methods for the
```

```

: TITLE OF INVENTION: Modification of Plant Lignin Content
: FILE REFERENCE: 11000.1003cad
: CURRENT APPLICATION NUMBER: US/09/615,192A
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 08/975,316
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: US 08/713,000
: PRIOR FILING DATE: 1996-09-11
: PRIOR APPLICATION NUMBER: US 09/169,789
: PRIOR FILING DATE: 1998-10-09
: NUMBER OF SEQ. ID NOS: 405
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 365
: LENGTH: 646
: TYPE: DNA
: ORGANISM: Pinus radiata
US-09-615-192A-365

Alignment Scores:
Pred. No.: 3,82e-56 Length: 646
Score: 624.00 Matches: 130
Percent Similarity: 75.12% Conservative: 21
Best Local Similarity: 64.68% Mismatches: 44
Query Match: 38,90% Indels: 6
DB: Gaps: 2

US-10-047-825-4 (1-313) x US-09-615-192A-365 (1-646)
QY      1  MetaLaserProthrLeuMetGlnCysLeuValAlaValSerLeu-----SerCys 18
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      44  ATGACTTCCTTTACAGAAATGAGGCGTCAGTGTGTGATGCTCTCTTTTTCGACC 103
QY      19  ValAlaHisAlaGlnLeuSerProThrPheArgAlaSerSerCysProAsnLeuGlnSer 38
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      104 GTTCCTTTGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 163
QY      39  IleValArgAlaAlaMetThrGlnAlaValAlaSerGluGlnArgMetGlyAlaSerLeu 58
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      164 ACGGTGAAGGCTGCAGGAAGAGCGGTGCCAAGAAAGAAAGGATGGGCGCATCATG 223
QY      59  LeuArgLeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeuLeuAsp 78
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      224 CTCGGCTTCCTTTCATGATGTTTGTGCAATGGTTGGAGGTGATGCTGATGTTGGAC 283
QY      79  AlaGly-----GlyGluysThrAlaGlyProAsnLeuAsnSerValArgGly 94
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      284 GACTCTTCGACTTAAGTGGGAAAGACTGCTCTTCCACACAAATTCGTTGAGGGGT 343
QY      95  PheGluValIleAspThrIleLysArgAsnValGluAlaIleCysProGlyValAlaSer 114
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      344 TTCGACGTCAATGACACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 403
QY      115  CysAlaAspIleLeuAlaLeuAlaValArgAspGlyThrAsnLeuLeuGlyProThr 134
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      404 TGCCTGACATTTGGCTATACGGCTAGAGATCTGTCTGTGAAATGGAAGAGACACACA 463
QY      135  TrpSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSer 154
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      464 TGGACAGGCTGTGGAAGAGAGAGACTCAGCAACTGCGCAGCCGCAAAACACC 523
QY      155  AsnProProProProProThrAlaSerLeuGlyThrIleLeuSerLeuPheGlyArgGlnGly 174
      ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      524 AACATTCCTCCGCTCCATTCCAATCTCAGTGTGATCTATCTTTTCAGACACAGGCG 583
QY      175  LeuSerProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCys 194
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      584 CTTTCAACCAAGATATGATGTTGCTATGAGTGTCAGTGACATATCATGTCAGAGCTGATGC 643
QY      195  Thr 195
      |||
Db      644 ACA 646
```



Wed May 7 14:10:04 2003

us-10-047-825-4\_1.rni

Page 13

Search completed: May 3, 2003, 13:16:30  
Job time : 55 secs



GenCore version 5.1.4-f5.4578  
Copyright (c) 1993 - 2003. Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 3, 2003, 12:58:59 ; Search time 108 Seconds

(without alignments)  
3424.513 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 1604  
Sequence: 1 MASPTLMQCLIVANVSLSCVA.....VGVLTGTAGQIRNCRVNS 313

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Dgapop 6.0	Dgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_model -DEV=xlib  
-O=/cgn2\_1/USPIO\_spool/US10047825/unat\_26042003\_112000\_7248/app\_query.fasta.1.455  
-DB=Published Applications\_NA -OFMT=fastap -SUFFIX=cmpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT-0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10047825.ecgn\_1\_1.84 -runat\_26042003\_112000\_7248  
-NCPU=6 -ICPU=3 -NO\_XLPEXT -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONLOG  
-DEV\_TIMECUT=120 -WARN\_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FEAPOP=6  
-FEAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA.\*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/FCI_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	808	50.4	996	US-09-938-842A-1711	Sequence 1711, Ap
2	801	49.9	1017	US-09-938-842A-60	Sequence 60, Appl
3	765.5	47.7	1008	US-09-938-842A-1714	Sequence 1714, Ap
4	753.5	47.0	1077	US-09-938-842A-1980	Sequence 1980, Ap

5	739.5	46.1	1011	US-09-938-842A-2545	Sequence 2545, Ap
6	725	45.7	966	US-09-938-842A-1273	Sequence 1273, Ap
7	717.5	44.7	1059	US-09-938-842A-2325	Sequence 2325, Ap
8	706	44.0	1041	US-09-938-842A-1632	Sequence 1632, Ap
9	699	43.6	1062	US-10-101-736-3	Sequence 3, Appl1
10	699	43.6	1062	US-09-834-659-3	Sequence 3, Appl1
11	699	43.6	1062	US-09-834-656-3	Sequence 1719, Appl
12	695.5	43.4	1011	US-09-938-842A-1719	Sequence 1457, Ap
13	692.5	43.2	1062	US-09-938-842A-1457	Sequence 2559, Ap
14	674.5	42.1	1011	US-09-938-842A-2559	Sequence 1518, Ap
15	658.5	41.1	987	US-09-938-842A-1518	Sequence 783, Appl
16	655.5	40.9	981	US-09-938-842A-783	Sequence 11, Appl
17	622.5	38.8	1281	US-10-027-559-111	Sequence 1566, Ap
18	619.5	38.6	960	US-09-938-842A-1596	Sequence 1340, Ap
19	616	38.4	981	US-09-938-842A-2367	Sequence 2367, Ap
20	610.5	38.1	942	US-09-938-842A-1498	Sequence 1498, Ap
21	605.5	37.7	966	US-09-938-842A-2693	Sequence 1193, Ap
22	596	37.2	945	US-09-938-842A-2693	Sequence 2693, Ap
23	586	36.5	993	US-09-938-842A-2656	Sequence 2656, Ap
24	584	36.4	1074	US-09-938-842A-2570	Sequence 2446, Ap
25	581	36.2	954	US-09-938-842A-2446	Sequence 1078, Ap
26	575	35.8	969	US-09-938-842A-1078	Sequence 682, Appl
27	575	35.8	990	US-09-938-842A-2329	Sequence 1549, Ap
28	569	35.5	978	US-09-938-842A-1682	Sequence 2079, Ap
29	564	35.2	978	US-09-938-842A-1549	Sequence 1548, Ap
30	564	35.2	960	US-09-938-842A-2079	Sequence 2181, Appl
31	562	35.0	990	US-09-938-842A-1548	Sequence 2209, Ap
32	560	34.9	990	US-09-938-842A-2209	Sequence 2184, Appl
33	545.5	34.0	1045	US-09-770-445-166	Sequence 50, Appl
34	540	33.7	1116	US-09-938-842A-2181	Sequence 858, Appl
35	527	32.9	990	US-09-938-842A-2184	Sequence 1458, Appl
36	522.5	32.6	990	US-09-938-842A-2184	Sequence 353, Appl
37	522.5	32.6	990	US-09-938-842A-2184	Sequence 889, Appl
38	508	31.7	1041	US-09-938-842A-858	Sequence 83, Appl
39	496	30.9	996	US-09-938-842A-1458	Sequence 361, Appl
40	485	30.2	1014	US-09-770-445-217	
41	477.5	29.8	992	US-09-770-445-353	
42	477	29.7	755	US-09-770-445-989	
43	433.5	27.0	824	US-09-894-635A-83	
44	415.5	25.9	1143	US-09-770-445-75	
45	396.5	24.7	1560	US-09-887-576-361	

#### ALIGNMENTS

RESULT 1  
US-09-938-842A-1711  
Sequence 1711, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kieps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SCRI300-3  
CURRENT FILING DATE: US/09/938, 842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227, 866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264, 647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300, 111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 1711  
LENGTH: 996  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1711  
Alignment Scores:

Pred. No.: 4.43e-88 Length: 996  
 Score: 808.00 Matches: 159  
 Percent Similarity: 66.25% Conservative: 51  
 Best local Similarity: 50.16% Mismatches: 95  
 Query Match: 50.37% Indels: 12  
 DB: 9 Gaps: 3

US-10-047-825-4 (1-313) x US-09-938-842A-1711 (1-996)

```

OY 9 CysLeuValAlaValSerLeuLeuSerCysValAlaAlaGlnLeuSerProThrPhe 28.
DB 43 TGCCTTGTCCCTCTCTCTCTCTGAGCAAGAGCTATGAGCGCAACTTTCCGGCTTAT 102
OY 29 TyrAlaSerSerCysProAsnLeuGlnSerIleValAlaAlaMethrGlnAlaVal 48
DB 103 TACGCCCACTTCTCCCAAGCAAGTCAATGAGTGGTGAATCACTTCTACCTAAAGCTGT 162
OY 49 AlaSerGlnArgMetGlyAlaSerLeuArgLeuPhePheHisAspCysPheVal 68
DB 163 GCTAGAGAGACAGCTATGCGTCTCTCTTGTGAGACTTATTTCCACGACTGTTCGTT 222
OY 69 GlnGlyCysAspGlySerIleLeuLeuAspAlaGly-----GluYsthr 84
DB 223 CAGGTTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 282
OY 85 AlaGlyProAsnLeuAsnSerValArgGlyPheGlyValIleAspThrIleLeuValArg 104
DB 283 TCMAACCTTAACCAACCAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 342
OY 105 ValGlnAlaAlaCysProGlyValAlaSerCysAlaAspIleLeuAlaLeuAlaAlaArg 124
DB 343 CTGGAAGAACAAATGCTCTGGAAGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 402
OY 125 AspGlyThrAsnLeuGlnGlyProThrTrpSerValProLeuGlyArgArgAspSer 144
DB 403 GACTCCTCTGTCTTACCGGTGAGCAAGTGGTGTCTTCTCTCTCTCTCTCTCTCTCTCT 462
OY 145 ThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProProThrAlaSerLeuGly 164
DB 463 AGAAGTGCACACCTGACATCAATCAACACACATCTCTGACCAACCAACACATTTTCCAG 522
OY 165 ThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMethrIleAlaSer 184
DB 523 AACCATTCTATCTAAGTTTAACCTGCAAGACTTGATTCATCACTGCTGCTGCTCTCTCC 582
OY 185 GlyAlaHisThrIleGlyGlnAlaArgCysThrPheArgGlyArgIleGly----- 203
DB 583 GGTAGTCACACAAATCGGATCTCTCGAGATGCAAGGTTTTCAGACAGAGGTTGTCAACAC 642
OY 204 -----AspThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGln 217
DB 643 TCCGGAACGGCAGCTCCAGCATGACATCTGGAACATCTCTCTCTCTCTCTCTCTCTCT 702
OY 218 ThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGlnThrProVal 237
DB 703 CGGTGTCCAAATATCCGGGGGAGACCAAGATTCTCTCGGTCTAGACATCATCAACGCG 762
OY 238 ArgPheAspThrAlaIlePheThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAsp 257
DB 763 AGCTTCGACAAACACTTCTTCAAGACTGTGATGAGAACAAGGCTGTGGAAGCTTGCAG 822
OY 258 GlnGlnLeuPheAsnGlnGlySerGlnAsp---AlaLeuValArgGlnIleYSerAlaSer 276
DB 823 CAAGTTCTGTTCACAGATTAACGAGAATTCGAGAGCTTGTGAAGAATATGCAAGAT 882
OY 277 AlaSerLeuPheAsnAlaAspPheValAlaAlaMethrIleArgMetGlyAsnValGlyVal 296
DB 883 CAAGAGAGTTTGTGACAAATTTGCGGAATTCATGATGATGGAATGGAATATCTCTCC 942
OY 297 LeuThrGlyThrAlaGlnIleArgArgAsnCysArgValAlaAsnSer 313
DB 943 TTGACAGGTTCTAGTGGGAAATCAGAGAAATTTGCGAAGATTTAACTCT 993

```

# RESULT 2

US-09-938-842A-60

; Sequence 60, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Krepes, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SCRIPT300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 60

; LENGTH: 1017

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-60

## Alignment Scores:

Pred. No.: 3.2e-87 Length: 1017  
 Score: 801.00 Matches: 156  
 Percent Similarity: 67.22% Conservative: 47  
 Best local Similarity: 51.66% Mismatches: 87  
 Query Match: 49.94% Indels: 12  
 DB: 9 Gaps: 3

US-10-047-825-4 (1-313) x US-09-938-842A-60 (1-1017)

```

OY 24 LeuSerProThrPheThrAlaSerSerCysProAsnLeuGlnSerIleValArgAlaAla 43
DB 106 CTCCTCCAGAGTTCTTCAACAACTGCTGCTGAGCCGAGAGATTCGAGTCACTT 165
OY 44 MethrGlnAlaValAlaSerGlnArgGlnArgMetGlyAlaSerLeuLeuArgLeuPhe 63
DB 166 GTACCAAGCTGTTCGAGAGAGAGACTGATGCTGCTCTCTCTCTCTCTCTCTCTCT 225
OY 64 HisAspCysPheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGly----- 81
DB 226 CACGATTGTTTCGTTTCAGGCTGTGATGATCGTCTCTTACACACCACTGGAGATATA 285
OY 82 -----GluYsthrAlaGlyProAsnLeuAsnSerValArgGlyPheGlyValIleAsp 99
DB 286 GTTACTGAGAAAGACTTAAACCCGAAACAGCATCGGCTCGCGTTTGAAGTTGTGAC 345
OY 100 ThrIleLeuArgAsnValGlnAlaAlaCysProGlyValAlaSerCysAlaAspIleLeu 119
DB 346 GAGATCAAGCGACATTAAGAGCAAGATCCCTTAACACTGTTCTGTCTGATGCTTA 405
OY 120 AlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGlyProThrPheSerValProLeu 139
DB 406 ACTTACCGGTAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 465
OY 140 GlyArgArgSerThrThrAlaSerAlaSerAlaAsnSerAsnProProPro 159
DB 466 GGAAGAGAGATTTCGACAAAGTTCGAGAGCTTGAAGATCAACAAACAAATTCCTGAC 525
OY 160 ThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAsp 179
DB 526 AACAACTTTCACAAACAAATTTGACAGATTTTAACAAACAAAGTTTCATCTCAC 585
OY 180 MethrIleLeuSerGlyAlaIleIleIleIleIleIleIleIleIleIleIleIleIle 199
DB 586 GTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 645

```

```

QY 200 ArgIleTyrGly-----AspThrAspIleAsnAlaSerPheAla 212
    |||:|||||
Db 646 AGACTGTACAAACCAATCCGGAAGCAAGTCCGACAGAACTTAGCAATCTCAAGCT 705
QY 213 AlaLeuArgGlnGlnThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAsp 222
    |||:|||||
Db 706 GCTAACTTGGCCCAAGAGTCCCTAGATCGGGTGGGACCAAGAACTGTACAGACTTGAC 765
QY 233 ValIleThrProValArgPheAspThrAlaTyrPheThrAsnLeuSerArgArgGly 252
    |||:|||||
Db 766 ATCAACAGCGCTGCAAAATTTGATTAACAGCTACTCAAGAACTTGATCGAAGCAATGGGA 825
QY 253 LeuPheHisSerAspGlnGlnLeuPheAsnGlyGlySerGlnAsp--AlaLeuValArg 271
    |||:|||||
Db 826 CTTGTGAATTCGACGAGGCTTGTCTCTCAAGCAATGACAAATCAAGAGCGCTCGTCAAG 885
QY 272 GlnTyrSerAlaSerAlaSerLeuPheAsnAlaAspPheValAlaAlaMetIleArgMet 291
    |||:|||||
Db 886 AAGATATCGACAGATCAAGAAAGACTTCTCGAGCAGTTCGCGAGTCCGAGATGATCAAGATG 945
QY 292 GlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsnTyrArgValVal 311
    |||:|||||
Db 946 GGGATATCTCTCCCTTGACAGGTTGCGATGCGATCGAAGAAATTCAGAGAAATTT 1005
QY 312 AsnSer 313
    |||:|||||
Db 1006 AATAAC 1011

RESULT 3
US-09-938-842A-1714
; Sequence 1714, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1714
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1714

Alignment Scores:
Pred. No.: 6,14e-83 Length: 1008
Score: 765.50 Matches: 153
Percent Similarity: 65.00% Conservative: 55
Best Local Similarity: 47.81% Mismatches: 99
Query Match: 47.72% Indels: 13
Caps: 3

US-10-047-825-4 (1-313) x US-09-938-842A-1714 (1-1008)
QY 6 LeuMetGlnCysLeuValAlaValSerLeuLeuSerCysValAlaHisAlaGlnLeuSer 25
    |||:|||||
Db 40 AATATCACTCTTAATGTCATGTTCTTCATATTGGAGCATCATCTCGCGCACTAATAC 99
QY 26 ProThrPheTyrAlaSerSerCysProAsnLeuGlnSerIleValArgAlaAlaMetThr 45
    |||:|||||
Db 100 GCACATTTTACTGTGGAGCTTGTCCACAGCCCTCTGCATCGTAGCGAGCACTATTACG 159

```

```

QY 46 GlnAlaValAlaSerGlnGlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAsp 65
    |||:|||||
Db 160 CAAGCTCTTCATTCGATATACAGAAATCGAGGACCGCTCATCGGCTTCATTTTCATAC 219
QY 66 CysPheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGlyGly----- 81
    |||:|||||
Db 220 TGTTTTGTAAAGAGATGATCATCATGATCTGTGACGACACATGGAAGATCCAGAGC 279
QY 82 GlnTyrThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlnValIleAspThrIle 101
    |||:|||||
Db 280 GGAAGAACGCTGGTCCGATGATTAACAGCTAGAGGATTAACGTTGTATATATTC 339
QY 102 LysArgAsnValGlnAlaAlaCysProGlyValAlaSerCysAlaAspIleLeuAlaLeu 121
    |||:|||||
Db 340 AAGACTGCGCTCGAAGAAAGCTTGCCCTGGTGTGTCTCTGCTGTGCTGAGGTTTACCCCT 399
QY 122 AlaAlaArgAspGlyThrAsnLeuGlnGlyProThrTTPSerValProLeuGlyArg 141
    |||:|||||
Db 400 GCCTCTGAGGCTTCTGTCTTGTGGCAGAGGCGCATCATGACTGTATTATAGGAAGA 459
QY 142 ArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThrAla 161
    |||:|||||
Db 460 AGAGATAGTCTACAGGCCAACCTCGCGGTCGAATTCGTCATTCTTCCATTCGAA 519
QY 162 SerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThr 181
    |||:|||||
Db 520 AGCCTAAGCAACATCATTTAATTTTGGCTGCGGCTTAATACAAACAGCATCGTGA 579
QY 182 AlaLeuSerGlyAlaHisThrIleGlnAlaArgCysThrThrPheArgGlyArgIle 201
    |||:|||||
Db 580 GCCTTATCTGTGGCGCATCGTTCGGCGCTGCTGATGAGATTTTCAACACAGACTA 639
QY 202 TyrGly-----AspThrAspIleAsnAlaSerPheAlaLeu 214
    |||:|||||
Db 640 TTTAACCTTACGGGGACAGAAATCCGATCCCACTTAATCAACGCTTGTGAGCACT 699
QY 215 ArgGlnGlnThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGln 234
    |||:|||||
Db 700 CTTCACAGAGTATGCTCTCAAAAGGATGCGCATCAACGATCAACCAATCTGATTGAGC 759
QY 235 ThrProValArgPheAspThrAlaTyrPheThrAsnLeuSerArgArgGlyLeuPhe 254
    |||:|||||
Db 760 ACACCGATGCGTTCGATTAACAATTCTGCCAACCTTCAGACGACAGCAGTACTTCTT 819
QY 255 HisSerAspGlnGlnLeuPheAsn-----GlyGlySerGlnAspAlaLeuValArgGln 272
    |||:|||||
Db 820 CAGTCAGACCAAGAGCTGTTCTCTACACCGGTTATCCACCATGCGCATTTGTACTTCG 879
QY 273 TyrSerAlaSerAlaSerLeuPheAsnAlaAspPheValAlaAlaMetIleArgMetGly 292
    |||:|||||
Db 880 TTGCAAGATTAACCAAGACTGTTTTCAGGCCCTTGACAGCTCATGATCAATATGAGGG 939
QY 293 AsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsnCysArgValAlaAsn 312
    |||:|||||
Db 940 AATATTATGTCCTTGACTGGAGTATGAGAGATTAGGCTTAGACTGTAAAGAGGTATAT 999

RESULT 4
US-09-938-842A-1980
; Sequence 1980, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647

```

; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 1980  
 ; LENGTH: 1077  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-1980

## Alignment Scores:

Pred. No.:	191e-81	Length:	1077
Score:	733.50	Matches:	149
Percent Similarity:	65.31%	Conservative:	60
Best Local Similarity:	46.56%	Mismatches:	98
Query Match:	46.98%	Indels:	13
DB:	9	Gaps:	3

US-10-047-825-4 (1-313) x US-09-938-842A-1980 (1-1077)

QY 6 LeuMetGlnCysLeuValAlaSerLeuSerCysValAlaHisAlaGlnLeuSer 25  
 Db 43 ATCATGAGCCTTATGTTTATGCTTTCATGTTGGACATCATCTCGCAGTTAAAC 102  
 QY 26 ProThrPheTyrAlaSerSerCysProAsnLeuGlnSerIleValAlaAlaMetThr 45  
 Db 103 GCACGCTTTTACTACGAGCAGCATCCCTAACGCTCTGCGCATGCTGAGCAGCTATTCAG 162  
 QY 46 GlnAlaValAlaSerGlnArgMetGlyAlaSerLeuLeuArgLeuPheHisAsp 65  
 Db 163 CAGGCTTTCAATCCATGCAAGATGAGAGCGCTTATCGCTTATTTTCAGCAGC 222  
 QY 66 CysPheValGlnGlyCysAspGlySerIleLeuLeuAsp-----AlaGlyGly 81  
 Db 223 TGTGTTGTAATGAGTGGCATGGGCTGCTGCTTGCAGACACTTCAGCATTCAGACC 282  
 QY 82 GlnLysThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlnValIleAspThrIle 101  
 Db 283 GAGAGACAGCGCTCTCCATGCAATGCAACTCACTGAGAGATTCAATGTTGATGATGATC 342  
 QY 102 LysArgAsnValGlnAlaHisProGlyValAlaSerCysAlaAspIleLeuAlaLeu 121  
 Db 343 AAGACAGCCCTCGAAGATGCTTCCGGCATGTTGCTGCTGACATTTTACCTCTT 402  
 QY 122 AlaAlaArgAspGlyThrAsnLeuGlnGlyGlyProThrTrpSerValProLeuGlyArg 141  
 Db 403 GCCTCAGAGGCCCTGCTGCTTGGCAGAGGACCTTCATGACATGCTATTATAGACAA 462  
 QY 142 ArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThrAla 161  
 Db 463 AGAGATGGTCTCACCCCAACTTGTCTGGAGCCAAATTCCTCTCCCTCCGAA 522  
 QY 162 SerLeuGlyThrLeuLeuSerLeuPheGlyArgGlnGlyLeuSerProArgSpMetThr 181  
 Db 523 GGCCTTAACACATCACAATTAATTTGAGCTGTCGGCTAAAGACAACCATGATGTA 582  
 QY 182 AlaLeuSerGlyAlaHisThrIleGlyAlaAlaArgCysThrThrPheArgGlyIleArg 201  
 Db 583 TCCTTGTCGAGAGCCATACGTTGGGCGTGTGATGCTAAGCTTCAACAATGAGACTA 642  
 QY 202 TyrGly-----AspThrAspIleAsnAlaSerPheAlaAlaLeu 214  
 Db 643 TTCACCTTCAAGCGAGCAAAACCCCGACCTGCACTGCAACACTTTCAGCAGT 702  
 QY 215 ArgGlnGlnThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspAlaGln 234  
 Db 703 CTTCACACAGCTATGCTCTAAACGAGCAGCAATACAGGATACCAATCTCATCTGAGC 762  
 QY 235 ThrProValArgPheAspThrAlaTyrPheThrAsnLeuSerArgArgGlyLeuPhe 254  
 Db 763 ACACCGATGCGTTCGATTAACAATTAATCTCAGAACCTTCAAGATTAACAATGGCTTCTC 822  
 QY 255 HisSerAspGlnGlnLeuPheAsn-----GlyGlySerGlnAspAlaLeuValArgGln 272

Db 823 CAGTCAGACAGCAACTTCTCTCAACACCGGTTACACACCGTCCGATGTTAATTC 882  
 QY 273 TyrSerAlaSerAlaSerLeuPheAsnAlaAspPheAlaAlaMetIleArgMetGly 292  
 Db 883 TTTCAGATTAACAGACCCCTGTTTGGAGCGCTTGTTCAGTATGATCAAGATGGG 942  
 QY 293 AsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsnCysArgValAlaAsn 312  
 Db 943 AACATTACTCCCTTGACGTGAGTAGTGGAGAGATTAGACAGACTGTAGGTGTTAAT 1002

## RESULT 5

US-09-938-842A-2545  
 ; Sequence 2545, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE  
 ; FILE REFERENCE: SRIPI300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 2545  
 ; LENGTH: 1011  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-2545

## Alignment Scores:

Pred. No.:	8.54e-80	Length:	1011
Score:	739.50	Matches:	153
Percent Similarity:	61.47%	Conservative:	48
Best Local Similarity:	46.79%	Mismatches:	101
Query Match:	46.10%	Indels:	25
DB:	9	Gaps:	4

US-10-047-825-4 (1-313) x US-09-938-842A-2545 (1-1011)

QY 11 ValAlaValSerLeuLeuSerCysValAlaHisAla----- 22  
 Db 19 ATCCTTATGAGCTCTCTCGCTCATGATTTTCTCTTTGTTGTGTTCCAAAGCC 78  
 QY 23 -----GlnLeuSerProThrPheTyrAlaSerCysProAsnLeuGln 37  
 Db 79 TATGGAAGTGGGCTATCTCTTCTTCAGTTTATGACCAATCGTGTCTTAAAGCTCA 138  
 QY 38 SerIleValArgAlaAlaMetThrGlnAlaValAlaSerGlnArgMetGlyAlaSer 57  
 Db 139 GAGATTGTGTCATGCTCATTTGCTTAAGCAATTCGAACATATCTCGCATGCTCC 198  
 QY 58 LeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeuLeu 77  
 Db 199 TTGCTCAGACTCCATTTCACGAGCTGTTGCTCAAGGATGATGATGCTCATATGTTG 258  
 QY 78 AspAlaGlyGly-----GlnLysThrAlaGlyProAsnLeuAsnSerValArg 93  
 Db 259 GACAGAGGTGAGAACCATTAATCAGGAGAGAAACGATCAAAACCTTAACAGAACTCAGCCGT 318  
 QY 94 GlyPheGlnValIleAspThrIleLysArgAsnValGlnAlaHisProGlyValVal 113  
 Db 319 GGTTCGATCATCGAGAGAAATCAACATGCTTCAACAAGAGTGTCTGAAACAGTT 378  
 QY 114 SerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGlyPro 133

```

|||||
Db 379 TCTGGCGTGAATCTTGGCTGACCGCTAGACACATCACTGTATTAGGGGTGACCG 438
|||
Qy 134 ThrTrpSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsn 133
|||
Db 439 AGCTGGAGAACTCTAGGAAGAGACGCGAGACACAGCTTGATGGTTCACAC 438
|||
Qy 154 SerAsnProPropProThrAlaSerLeuGlyThrLeuLeuSerLeuPheGlyArgGln 173
|||
Db 499 AACACATCTCTGCTCCAAACACACTTCCAAACCTCTCACTAAGTTCAAGCTCA 558
|||
Qy 174 GlyLeuSerProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyAlaArg 193
|||
Db 559 GGCCTGATGTCGTGCAATCTGCTCCCTCGAGAGATGCACACCAATGGAACTGCAG 618
|||
Qy 194 CysThrThrPheArgGlyArgIleTyrgly-----AspThrAsp 206
|||
Db 619 TGACAAAGTTCCGACAGAGTTATACAAATCCGCAACGGAAGCCTGATATGACT 678
|||
Qy 207 IleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyAspGly 226
|||
Db 679 CTTAGCCAAATCTGCAACACTATTCGTACAGATGTCGAGATCCGAGGTGACAC 738
|||
Qy 227 AsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyrrPheThrAsn 246
|||
Db 739 ACCCTATTCTTCCTCGACTTCGCGACACCGTCAAGTTGACAAACCACTACTTCAAGAC 798
|||
Qy 247 LeuLeuSerArgArgGlyLeuPheHisSerAspGlnGlnLeuPheAsnGlyGlySerGln 266
|||
Db 799 CTGATATGTACAAAGGCTCATTCAGCTCTGATGAGATGCTGTTCACGAGAACCAAC 858
|||
Qy 267 Asp---AlaLeuValArgGlnTyrrSerAlaSerAlaSerLeuPheAsnAlaAspPheVal 285
|||
Db 859 TCCAAGAGACTGCGTGGAGCTATACGCTGACATCAAGAGCCCTTTTGACGAGTTGCT 918
|||
Qy 286 AlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArg 305
|||
Db 919 AAATCGATGGAGAGATGAGAAATATCTCTCCGTGACAGCGCGCAAGGAGAGATCAG 978
|||
Qy 306 ArgAsnCysArgValAlaAsn 312
|||
Db 979 CGTATCTGTCGAGGCTTAAC 999
|||

RESULT 6
US-09-938-842A-1273
: Sequence 1273 Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPI300-3
: CURRENT APPLICATION NUMBER: US/09/938,842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 1273
: LENGTH: 966
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-1273

```

Alignment Scores: 4.51e-78 Length: 966  
 Pred. No.: 725.00 Matches: 162  
 Score:

```

Percent Similarity: 58.77% Conservative: 29
Best Local Similarity: 49.85% Mismatches: 118
Query Match: 45.20% Indels: 16
DB: 9 Gaps: 6

US-10-047-825-4 (1-313) x US-09-938-842A-1273 (1-966)
Qy 1 MetAlaSerProThrIleuMetGlnCysLeuValAlaValSerLeuLeuSerCysValAla 20
|||
Db 1 ATGGAGGACATGCTTGCTGCTTACTGTTTCACTCTTTGTATGCTTGTGTGACAGGGTT--- 57
|||
Qy 21 HisAlaGlnLeuSerProThrPheTyrrAlaSerSerCysProAsnLeuGlnSerIleVal 40
|||
Db 58 AGGCGACACTAAGCCCTACACTTATATGCTTAATGCTGCCCAATCTGCTCAAAATTTGTC 117
|||
Qy 41 ArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGlyAlaSerLeuLeuArg 60
|||
Db 118 CGTAACAAAGTTGCCATCGCCCTGAAGCCGAGATACGATGCTGCTTCATTCGT 177
|||
Qy 61 LeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeuLeuAsp---Ala 79
|||
Db 178 CTTGATTTCCACGACACTGCTTGTATATGCTGTGATGCTGCTTATTTGTGTGATGACT 237
|||
Qy 80 GlyGlyGlnTyrrThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlnValIleAsp 99
|||
Db 238 GACAGGAGAAACTCGCCGATCCCAACATTAATCTGCTAGAGGATTTGAGATATTTGAT 297
|||
Qy 100 ThrIleTyrrArgAsnValGlnAlaAlaCysProGlyValAlaSerCysAlaAspIleLeu 119
|||
Db 298 ACAATCAAAAGCCGCTGTGTAAGAAAGATGCTCTGGTGTTGTTCTGTGATGATTCCTC 357
|||
Qy 120 AlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGlyProThrTrpSerValProLeu 139
|||
Db 358 ACTTGGCCGCTGCGTACCTCCCTGTTGATGAGTGAAGGCGCTGGGTGAGATGGCATTA 417
|||
Qy 140 GlyArgArgAspSerThrThrAlaSerAlaSerIleuAlaAsnSerAspProProPro 159
|||
Db 418 GGAAGAAAGATGATGTTGGTGGCAATGACAGACATGCAAC---AATCTACATCTCTCT 474
|||
Qy 160 ThrAlaSerLeuGlyThrLeuLeuSerLeuPheGlyArgGlnGlyLeuSerProArgAsp 179
|||
Db 475 TTTGAACCTTTAGCGCCATTTATCGCAAAATTTGACCGTTAACTTAACATCACCGAC 534
|||
Qy 180 MetThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArgGly 199
|||
Db 535 GTCTGATGCTTTATACAGGACTCACACTTTTGACAAAGGAAAGTGTGCTCTTACGACAC 594
|||
Qy 200 ArgIleTyrgly-----AspThrAspIleAsnAlaSerPheAla 212
|||
Db 595 CGTCTGTTCACTTACCGGCTTGGGAAATCCCGACGCAACACTTGACATCACTCTTG 654
|||
Qy 213 AlaLeuArgGlnGlnThrCysProArgSerGlyGlyAsnLeuValAsnLeuAlaProIleAsp 232
|||
Db 655 TCTAATCTCAAAACAGTTTGTCCGCTCGAGGAAATAGTAACTAATACAGCACTCTTGAC 714
|||
Qy 233 ValGlnThrProValArgPheAspThrAlaTyrrPheThrAsnLeuLeuSerArgArgGly 252
|||
Db 715 AGGAGACATPACGAGACACTTCGACAAACAATTAATTTCAAGAACTGCTTGAAGAAAGAGT 774
|||
Qy 253 LeuPheHisSerAspGlnGlnLeuPheAsnGly-----GlySerGlnAspAla 268
|||
Db 775 CTTTGGAGTTCGATCAAGTTCTGTCTTCGAGTGAAGTCAACCGTGAACCAACAAAGAA 834
|||
Qy 269 LeuValArgGlnTyrrSerAlaSerAlaSerLeuPheAsnAlaAspPheValAlaAlaMet 288
|||
Db 835 CTAGTGAAGGCTTATAGTCGAGCCAGCTTGTTCCTCAAGGACTTCAATATGTCGATG 894
|||
Qy 289 IleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsnCys 308
|||
Db 895 ATCAGAAATGGAAACATTTGCG-----AATGAGCTAGTGGGAGGTTAGGAAACAATGCG 948
|||
Qy 309 ArgValAlaAsnSer 313
|||

```

[illegible]

Db 514 TTCACCTCCCAACAACCTTAAGGCCAGTTTTCGAATAATGTGGACCTGACCGCCTCTGAT 573

Qy 180 MethrhalaleuSerglyAlaHisThrTlleGlyAlaIarGcYsThrPheArgGly 199  
:::|||||:::|||||:::|||||:::|||||

Db 574 CTCGGTGAACCTCTCGTGTGTACACACATTTGTGTAACCAACGAGTCCAAATTTATATGAT 633

Qy 200 ArgIleTyArgLy-----AspThrAspIleAsnAlaSerPheAla 212  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 634 AGACATATCAACACTTATAGCAACACTGGTTTACCAGACCCCTACTCTTAATACTACATACCTC 693

Qy 213 AlaIeArgInglInThrCysProArgSergIylGlyAspGlyAsnAlaIaProIleAsp 232  
:::|||||:::|||||:::|||||:::|||||

Db 694 CAGACACTTCGTCGGCAATGTCCCGCAATGTGAACCGACCGCTTAATGATTTGGAT 753

Qy 233 ValGlnThrProValArgPheAspThrAlaTyPheThrAsnLeuLeuSerArgGly 252  
:::|||||:::|||||:::|||||:::|||||

Db 754 CTACGCACACACGAGCGTTTGTGACAAATAATCATATGATCATGTGAAGAGACCTTAAGGGA 813

Qy 253 LeupheHisSerAspGlnGluLeuPhe-----AsnGlyGlySergInAspAlaLeu 269  
|||:::|||||:::|||||:::|||||:::|||||

Db 814 CTTATCCAGACGGAGCAAGAGTGTGTTTCCAGCCCTTAATGCGCACTGACACAAATCCCTTG 873

Qy 270 ValArgGlnTySerAlaSerAlaSerLeupheAsnAlaAspPheAlaIaIaMetIle 289  
|||||:::|||||:::|||||:::|||||

Db 874 GTGAGAGATATACGTCGATGCGACATCAAAATTTCTCAATGCTTTTGTGGAAGCAATGAT 933

Qy 290 ArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyInIleArgArgAsnCysArg 309  
:::|||||:::|||||:::|||||:::|||||

Db 934 AGAATGGGAAACATTTACACTCTCTCACTGAACTCAAGGACAAATCAGACAGAAATTTGATG 993

Qy 310 ValValAsnSer 313  
|||||:::|||||

Db 994 GTGGTTAACTCC 1005

RESULT 8  
US-09-938-842A-1632  
; Sequence 1632, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938, 842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227, 866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264, 647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300, 111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1632  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1632

Alignment Scores:  
Pred. No.: 9.94e-76 Length: 1041  
Score: 706.00 Matches: 147  
Percent Similarity: 63.79% Conservative: 45  
Best Local Similarity: 48.84% Mismatches: 95  
Query Match: 44.01% Indels: 14  
DB: 9 Gaps: 4

US-10-047-825-4 (1-313) x US-09-938-842A-1632 (1-1041)

Qy 22 AlaGlnLeuSerProThrPheTyrrAlaSerSerCysProAsnLeuGlnSerIleValArg 41  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||



```

Db 124 AGCAATCTCAACATGATCTATGACCGGCTTGTCCGGCTTCAACGATTTAA 183
QY 42 ALAlaMetThrGlnAlaValAlaSerGluGlnArgMetGlyAlaSerLeuAlaGlu 61
Db 184 TCCGAGATTGGAGAGCTTTTAAAGATGATTCGAAATCGCTGATCTTTTCGACTC 243
QY 62 PhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeuLeu----- 77
Db 244 CATTTCCACGATTTTGTTCATGAATGATGATGATCTATCTTCTGAAACGATTCAG 303
QY 78 AspAlaGlyGlyGluValGlyThrAlaGlyProAsnLeuAsnSerValArgGlyPheGluVal 97
Db 304 GATTTCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
QY 98 IleAspThrIleValArgAsnValGlnAlaAlaCysProGlyValValSerCysAlaAsp 117
Db 364 ATTAAGACATTTAATCTGATATGATGAAAGTCTTGTCCCTTAACAGTTATAGCGCTGAC 423
QY 118 IleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuGlyGlyProThrTrpSerVal 137
Db 424 ATAGTGTCTCGCGGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
QY 138 ProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnPro 157
Db 484 CCTTTGGAGCAGAGGAGCTCATTTAAGCGCGAGTGAGCAAGCGCGAGATACAAATCTGCCA 543
QY 158 ProProThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerPro 177
Db 544 TCTCCGTTTGGAGCGGTGGAGAGATATACAGCCAGAGTTCGACCCCTTGGACCTGCACCTC 603
QY 178 ArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrPhe 197
Db 604 AAGAGCGTGTGGTCTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
QY 198 ArgGlyArgIleTyrgly-----AspThrAspIleAsnAlaSer 210
Db 664 AAGCATAGACTCTTCACCTTCACAGGCGTCAAGCCAGCCGCTTCCAAACCTTGGCGCTTCC 723
QY 211 PheAlaAlaLeuArg-----GlnGlnThrCysProArg---SerGlyGlyAspGlyAsn 227
Db 724 TCAGCAGCTTCTCTTAAGCTTAAGAGACAGTGTCTTAACGTGAGCTCTCTCAGACTCTTAAG 783
QY 228 LeuAlaProIleAspValGlnThrProValArgPheAspThrAlaArgPheThrAsnLeu 247
Db 784 CTCGCTGCTTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
QY 248 LeuSerArgArgGlyLeuPheHisSerAspGlnGluLeuPheAsnGlyGlySerGlnAsp 267
Db 844 ATGACACACATAGGAGCTGTGGATTTGATCAAAACCTTAATGACAGATCCTACGCGTGC 903
QY 268 AlaLeuValArgGlnTySerAlaSerAlaSerLeuPheAsnAlaAspPheValAlaAla 287
Db 904 GCCCTTGGAAGTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 963
QY 288 MetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsn 307
Db 964 ATGGTTAAATAGCGGAATATGAGAGATTTGACCGGAGAGATGATGAGATATTGAGAGAAA 1023
QY 308 Cys 308
Db 1024 TGT 1026

```

```

; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 3
; SEQ ID NO 3
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Amoracia rusticana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1062)
; PUBLICATION INFORMATION:
; AUTHORS: Kazuhito FUJIIYAMA et al.
; TITLE: Structure of the horseradish peroxidase isozyme C genes
; JOURNAL: European Journal of Biochemistry
; VOLUME: 173
; PAGES: 681-687
; DATE: 1998
; US-10-101-736-3

Alignment Scores:
Pred. No.: 7,17e-75 Length: 1062
Score: 699.00 Matches: 153
Percent Similarity: 58.13% Conservative: 40
Best Local Similarity: 46.08% Mismatches: 119
Query Match: 43.58% Indels: 20
DB: Gaps: 5

US-10-047-825-4 (1-313) x US-10-101-736-3 (1-1062)
QY 2 AlSerProThrIleuMetGlnCysLeuValAlaValSerLeuLeuSerCysValAlaHis 21
Db 13 TCTTCTTCACTTGTTCATCTGTATTAACCTTAATCCATTTGGATGCTTATCTTCAT 72
QY 22 -----AlaGlnLeuSerProThrPheTyAlaSerSerCysProAsnLeu 36
Db 73 GCTTCTTGTCTGATGCTCACTTACCCTTACCTTATGACATTCATGCTTAATGTC 132
QY 37 GlnSerIleValArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGlyAla 56
Db 133 TCAACATCTGACGGGATGATATTGCAATGACCTTAAGTACAGACCCCTGATTTGCCG 192
QY 57 SerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeu 76
Db 193 AGCATCCTTCTGCTTCACTTCCAGCAGCTGCTTTGTAATGTTGTGACGATGATCTTG 252
QY 77 LeuAsp-----AlaGlyGlyGluValThrAlaGlyProAsnLeuAsnSerVal 92
Db 253 TTGACACACACATCATTTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
QY 93 ArgGlyPheGluValIleAspThrIleValArgAsnValGlnAlaAlaCysProGlyVal 112
Db 313 AGAGGATTTCCAGATGATGATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372
QY 113 ValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGly 132
Db 373 GTTTCAGGCCAGATTTGTCACCATTCAGCTCAACATCTGTCACTTTGGCGGAGAGT 432
QY 133 ProThrTrpSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAla 152
Db 433 CTTCTTGGAGAGTCTTCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
QY 153 AsnSerAsnProProProProThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArg 172
Db 493 AATGCAAAATCTTCAGAGTCTCATCTTCCACACTTCCACAACTTAAGACAGCTTACAAAT 552
QY 173 GlnGlyLeuSer---ProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGln 191
Db 553 GTTGGCCTCAACCGTCTTCTGATCTGCTGACTCTCCGGGGCCACACATTTGGTAA 612
QY 192 AlaArgCysThrThrPheArgGlyArgIleTyrgly-----Asp 204
Db 613 AATCAGTGTGCTTATTTATGACAGATTTATACAACTTACGACACACCGGTTTACCGCAT 672
QY 205 ThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyGly 224

```

```

RESULT 9
; Sequence 3, Application US/10101736
; Publication No. US20030041351A1
; GENERAL INFORMATION:
; APPLICANT: Yoshinaka KASUKABE, Izumi IHARA, Yoshinaka MAEKAWA (Toyoobo
; APPLICANT: Co., Ltd.); Randy Dale Allen (Texas Tech University)
; TITLE OF INVENTION: Cotton plants with improved cotton fiber characteristics and met
; FILE REFERENCE: for producing cotton fibers from these cotton plants
; CURRENT APPLICATION NUMBER: US/10/101,736

```

```

Db      673 CCACTCTCAACACTACTATCTCCAAACTCTCTGAGACTATGTCCTCCATAGTAT 732
      225 AspglyAsnleuAlaProIleAspValGlnThrProValArgPheAspThrAlaArgPhe 244
      733 CTAAAGCGCTTGGTGGATTTTGTATCTAGTCAGCCCAAGCATTTTGTACAAACAATCTAT 792
      245 ThrAsnleuSerArgArgGlyLeuPheHisSerAspGlnGluLeuPhe----- 261
      793 GTGAATCTCGAAGCAAAAAGAGCTTATCCAAAGGAGACCAAGAGTGTCTCTAGCCCC 852
      262 AsnGlyGlySerGlnAlaLeuValArgGlnIleThrSerAlaSerAlaSerLeuPheAsn 281
      853 AATGCCACTGACACAATCCCTTGTGAGATCATTTGCTAATAGCACACAACATTCCTTC 912
      282 AlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAla 301
      913 AATGCCGTTGTGGAGCGCATGATGATGAGATGGGAAACATTACACCTCTTACAGAACTCAA 972
      302 GlyGlnIleArgArgAsnCysArgValValAsnSer 313
      973 GGACAGATCAGGTTGATTTAGGTGAGTGAATCTCC 1008

```

## RESULT 10

```

; Sequence 3, Application US/09834659
; Publication No. US20030074697A1
; GENERAL INFORMATION:
; APPLICANT: Yoshinaka KASUKABE, Izumi IHARA, Yoshinaka MAEKAWA (Toyobo
; APPLICANT: Co., Ltd.); Randy Dale Allen (Texas Tech University)
; TITLE OF INVENTION: Cotton plants with improved cotton fiber characteristics and meth
; FILE REFERENCE: 204552015600
; CURRENT APPLICATION NUMBER: US/09/834,659
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US/09/347,669
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 3
; SEQ ID NO 3
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Armoreria rusticana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1062)
; PUBLICATION INFORMATION:
; AUTHORS: Kazuhito FUJIIYAMA et al.
; TITLE: Structure of the horseradish peroxidase isozyme C genes
; JOURNAL: European Journal of Biochemistry
; VOLUME: 173
; PAGES: 681-687
; DATE: 1988
; US-09-834-659-3

```

## Alignment Scores:

```

Pred. No.: 7,17e-75 Length: 1062
Score: 699.00 Matches: 153
Percent Similarity: 58.13% Conservative: 40
Best Local Similarity: 46.08% Mismatches: 119
Query Match: 43.58% Indels: 20
DB: 9 Gaps: 5

```

US-10-047-825-4 (1-313) x US-09-834-659-3 (1-1062)

```

      2 AlaSerProThrLeuMetGlnCysLeuValAlaValSerLeuLeuSerCysValAlaHis 21
      13 TCTTCTTCTACTTGTCTTCTGATATACCTTAATCCCATTTGATGCTATATCTTCTCAT 72
      22 -----AlaGlnLeuSerProThrPheTyrAlaSerSerCysProAsnLeu 36
      73 GCTTCTTCTGATGCTCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 132
      37 GlnSerIleValArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetCysVal 56

```

```

      133 TCTAATCATCGTAGCGGATACATATGTCATAGAGTGAAGATCAGACCCCTCGATTCGCCGC 192
      57 SerLeuLeuAlaGluPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeu 76
      193 AGCATCCCTTCCTTCTACCTTCCAGACGCTTCTTATGATGTGTGAGCGCATCATCTTG 252
      77 LeuAsp-----AlaGlyGlyLysThrAlaGlyProAsnLeuAsnSerVal 92
      253 TTAGACACACACATCATTTTCGAAAGAGAAAGATGCGTTTGGAACGCAATTCGCCCA 312
      93 ArgGlyPheGluValIleAspThrIleLysArgAsnValGluAlaAlaCysProGlyVal 112
      313 AGAGATTTCCAGATGATGATAGATGAATGAACCCCGGTGAGAGTGCATGCCCAAGAAC 372
      113 ValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGly 132
      373 GTTTCATCGCAGATTTGCTACCTTGCACCTTCCACACTTCCACAACTTAAGACACTTGA 432
      133 ProThrIlePheValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAla 152
      433 CTTCTTGGAGAGTTCCTTTGGGAGAGATAGATAGCTTACAGATTTCTGAGATCTTGCT 492
      153 AsnSerAsnProProProProThrAlaSerLeuGlyThrIleuIleSerLeuPheGlyArg 172
      493 AATGCAATCTTCCAGCTCCATTCCTTCCACACTTCCACAACTTAAGACACTTGA 552
      173 GlnGlyLeuSer-----ProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGln 191
      553 GTTGGCCCTCAACCGCTTCTTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
      192 AlaArgCysThrThrPheArgGlyArgIleGlyGly-----Asp 204
      613 AATCAGTGTGCTTATATATGACAGATTTATACAACTTCCAGCAACACCGGTTTACCCGAT 672
      205 ThrAspIleAsnAlaSerPheAlaAlaLeuArgGlnIleThrCysProArgSerGlyGly 224
      673 CCACTCTCAACACTACTTATCTCAAACTCTGCTGAGATCTCCCTCCATGATGAT 732
      225 AspglyAsnleuAlaProIleAspValGlnThrProValArgPheAspThrAlaArgPhe 244
      733 CTAAAGCGCTTGGTGGATTTTGTATCTAGTCAGCCCAAGCATTTTGTACAAACAATCTAT 792
      245 ThrAsnleuSerArgArgGlyLeuPheHisSerAspGlnGluLeuPhe----- 261
      793 GTGAATCTCGAAGCAAAAAGAGCTTATCCAAAGGAGACCAAGAGTGTCTCTAGCCCC 852
      262 AsnGlyGlySerGlnAlaLeuValArgGlnIleThrSerAlaSerAlaSerLeuPheAsn 281
      853 AATGCCACTGACACAATCCCTTGTGAGATCATTTGCTAATAGCACACAACATTCCTTC 912
      282 AlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAla 301
      913 AATGCCGTTGTGGAGCGCATGATGATGAGATGGGAAACATTACACCTCTTACAGAACTCAA 972
      302 GlyGlnIleArgArgAsnCysArgValValAsnSer 313
      973 GGACAGATCAGGTTGATTTAGGTGAGTGAATCTCC 1008

```

## RESULT 11

```

; Sequence 3, Application US/09834656
; Patent No. US20020049999A1
; GENERAL INFORMATION:
; APPLICANT: Yoshinaka KASUKABE, Izumi IHARA, Yoshinaka MAEKAWA (Toyobo Co., Ltd.);
; APPLICANT: Randy Dale Allen (Texas Tech University)
; TITLE OF INVENTION: Cotton plants with improved cotton fiber characteristics and m
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/834,656
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 3
; SEQ ID NO 3

```

: LENGTH: 1062  
 : TYPE: DNA  
 : ORGANISM: *Armoracia rusticana*  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: (1)...(1062)  
 : PUBLICATION INFORMATION:  
 : AUTHORS: Kazuhito FUJIYAMA et al.  
 : TITLE: Structure of the horseradish peroxidase isozyme C genes  
 : JOURNAL: European Journal of Biochemistry  
 : VOLUME: 173  
 : PAGES: 681-687  
 : DATE: 1988  
 : US-09-834-656-3

## Alignment Scores:

Pred. No.:	7.17e-75	Length:	1062
Score:	699.00	Matches:	153
Percent Similarity:	58.13%	Conservative:	40
Best Local Similarity:	46.08%	Mismatches:	119
Query Match:	43.58%	Indels:	20
DB:	10	Gaps:	5

US-10-047-825-4 (1-313) x US-09-834-656-3 (1-1062)

OY 2 AlaSerProThrLeuMetGlnCysLeuValAlaValSerLeuSerCysValAlaHis 21  
 DB 13 TCCTTCTTACTTGTTCACCTTGATTAACCTTAATCCCATGGATGCTTATCTTCAT 72  
 OY 22 -----AlaGlnLeuSerProThrPheTyraIaSerSerCysProAsnLeu 36  
 DB 73 GCTTCTTGTGTGATGCTCACTTACCCCTTACCTTCACGACCAATTCATGCTTAATGTC 132  
 OY 37 GlnSerIleValArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGlyAla 56  
 DB 133 TCCTACATCGTACGGGACTACTATGTCATAGATAGATAGACCTCGATTTGCGCG 192  
 OY 57 SerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeu 76  
 DB 193 AGCATCCTTCGCTTCTTCACTTCACACACTGCTTTGTAATGGTTGGAGCATTCATCTTG 252  
 OY 77 LeuAsp-----AlaGlyGlnLeuValGlnValGlnValProAsnLeuAsnSerVal 92  
 DB 253 TTAGACAAACACAAATCATTTTCGACAGAGAAAGATGCGTTGGAACGCAAACTCGGCA 312  
 OY 93 ArgGlyPheGlnValIleAspThrIleLysArgAsnValGlnAlaIleCysProGlyVal 112  
 DB 313 AGAGGATTTCCAGATGATGATGAATGAAGCCGCGGTGGAGAGTGCATGCCCAAGAAC 372  
 OY 113 ValSerCysAlaAspIleLeuAlaLeuAlaValArgAspGlyThrAsnLeuLeuGly 132  
 DB 373 GTTTCATGGCGCATTTGCTGCACATTCGACGTCAACATCGTCACTTTGGCGGAGAGT 432  
 OY 133 ProThrPheSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAla 152  
 DB 433 CCTTCTTGAGAGATTCCTTTGGCGAGAGATACCTTACCAAGATTTCTGTGATCTTCT 492  
 OY 153 AsnSerAspProProProProThrAlaSerLeuGlyThrIleLeuIleSerLeuPheGlyArg 172  
 DB 493 AATGCAAAATCTTCACGCTTCATTTCCACACTTCCACAACTTAAGAGACGTTTGAAT 552  
 OY 173 GlnGlyLeuSer---ProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGln 191  
 DB 553 GTTGGCGCTCAACCGTTCTTCTGATCTCGTTCGACATGTCGGGGGCGACACATTTGGTAA 612  
 OY 192 AlaArgCysThrThrPheArgGlyArgIleTyrgly-----Asp 204  
 DB 613 AATCATGTGCGTTATATATATGAGACAGATTTATCAACTTCAGACACACCGGTTTACCAG 672  
 OY 205 ThrAspIleAsnAlaSerPheAlaAlaLeuAlaValArgGlnGlnThrCysProArgSerGly 224  
 DB 673 CCTACTCTCAACACACTTATCTCCAAACTTTCGTGGGACTGTGTCCTCAATGTATAT 732

OY 225 AspGlyAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyrrPhe 244  
 DB 733 CTAAAGCGCTTGGTGGATTTGATCTTACGTACGCAACGATTTTGACAAACAAATACTAT 792  
 OY 245 ThrAsnLeuLeuSerArgArgGlyLeuPheHisSerAspGlnGluPhe----- 261  
 DB 793 GTGAATCTCGAAGAGCAAAAGACTTATCCAAAGCGACCAAGATGTTCTCTAGCCCC 852  
 OY 262 AsnGlySerGlnAspAlaLeuValArgGlnTyrrSerAlaSerAlaSerLeuPheAsn 281  
 DB 853 AATGCCACATGACCAATCCCTTGTGTGATGATCTTTGCTTAATAGCACAAACATCTTC 912  
 OY 282 AlaAspPheValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAla 301  
 DB 913 AATGCGTTTGTGAGGAGGATGATAGATGGAAACATTAACCTCTTACAGAACTCA 972  
 OY 302 GlyGlnIleArgArgAsnArgValAlaAsnSer 313  
 DB 973 GGACAGATCAGGTTGAATGTTAGGTTGGAAGTCC 1008

## RESULT 12

US-09-938-842A-1719

: Sequence 1719, Application US/09938842A  
 : Patent No. US20020160378A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Harper, Jeff  
 : APPLICANT: Kreps, Joel  
 : APPLICANT: Wang, Xun  
 : APPLICANT: Zhu, Tong  
 : TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 : FILE REFERENCE: SAME, AND METHODS OF USE  
 : CURRENT APPLICATION NUMBER: US/09/938, 842A  
 : PRIOR FILING DATE: 2001-08-24  
 : PRIOR APPLICATION NUMBER: US 60/227, 866  
 : PRIOR FILING DATE: 2000-08-24  
 : PRIOR APPLICATION NUMBER: US 60/264, 647  
 : PRIOR FILING DATE: 2001-01-16  
 : PRIOR APPLICATION NUMBER: US 60/300, 111  
 : PRIOR FILING DATE: 2001-06-22  
 : NUMBER OF SEQ ID NOS: 5379  
 : SEQ ID NO 1719  
 : LENGTH: 1011  
 : TYPE: DNA  
 : ORGANISM: *Arabidopsis thaliana*  
 : US-09-938-842A-1719

## Alignment Scores:

Pred. No.:	1.77e-74	Length:	1011
Score:	695.50	Matches:	153
Percent Similarity:	62.91%	Conservative:	37
Best Local Similarity:	50.66%	Mismatches:	95
Query Match:	43.36%	Indels:	17
DB:	9	Gaps:	4

US-10-047-825-4 (1-313) x US-09-938-842A-1719 (1-1011)

OY 28 PheTyraIaSerSerCysProAsnLeuGlnSerIleValArgAlaAlaMetThrGlnAla 47  
 DB 100 TTTCNCAAGAGATTTGCCATTTGGCAGAGAAATGCTGAACATTAACGAATGCT 159  
 OY 48 ValAlaSerGlnGlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAspCysPhe 67  
 DB 160 GTTCTCAAAAGATCCTGATGCGCGCTTCTTCTTCTGCTCAGTTCACGATTTGCTTC 219  
 OY 68 ValGlnGlyCysAspGlySerIleLeuLeuAspAlaGlyIle-----GlnLys 83  
 DB 220 GTCTGGGTTGTGATGCTGCTGCTTGGACACACATGAGATATATGTTCAAGTGA 279  
 OY 84 ThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlnValIleAspThrIleLysArg 103  
 DB 280 CAACCACTCAAAATCTGAAGTCTTTTACGTGATTTGAGTAATGATTCATTAAGTAC 339

```

QY 104 AsnValAlaAlaCysProGlyValValSerCysAlaAspIleuAlaAla 123
    ::::: |||||
Db 340 TCCTCGAAGAGCTTCTCTCACTCACTGTTCCGATATCCCTGCTGGCGCT 399
QY 124 ArgAspIleuThrAsnLeuLeuGlyGlyProThrTyrSerValProLeuGlyArgAsp 143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 400 CGTACCTCACTCTTCCGAGAGGCGCCATGCTGGAATTTTGGTAGGAGAAAGAC 459
QY 144 SerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThrThrAlaSerLeu 163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 460 TCATTAAGCGCAAGTTCCGCCGAGGAGAACCAATTCATTCACACCAATCTCTCTC 519
QY 164 GlyThrIleuIleuSerIleuPheGlyArgGlyGlyLeuSerProArgAspMetThrAlaLeu 183
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 520 GACAGCCTCATCATCATATTCACAGCAAGGCTCAACCAATTCATCATCATCGCTCTT 579
QY 184 SerGlyAlaHisThrIleGlyGlyAlaArgCysThrThrPheArgGlyArgIle----- 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 580 TCGGCTCTCATCATCATAGTAAAGCGAGATGCGTGAAGCTTCAGACGACGATTTCAA 639
QY 202 -----TyrGlyAspThr--AspIleAsnAlaSerPheAlaAla 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 640 CCAAACTGAGCAAAAGCTCTACGTCGATGAGTTCAGAAAGACAGCAAGTTCGCTGA 699
QY 214 LeuArgGlnIleThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspVal 233
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 700 GTCCTAGGCTCAGCAATTCGAAAGATTCAGCCGAGACAGCAACTCTCTCCACTGATATT 759
QY 234 GlnThrProValArgPheAspThrAlaTyrPheThrAsnLeuSerArgArgGlyLeu 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 760 AAGACACACCTTACTTCAATCAATCATTTATTCATCAATCTGTTAAAGAGAGAGGCTTG 819
QY 254 PheHisSerAspGlnLeuLeuPheAsnGlyGlySerGlnAspAlaLeu-----Val 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 820 TTGATTCGATTAACCTTCTGCTCGCAAGATCATGAAGAGAGATCTTCAGAAAGTT 879
QY 271 ArgGlnTyrSerAlaSerAlaSerLeuPheAsnAlaAspPheAlaAlaMetIleArg 290
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 880 TGGGAATATTCATTAATCAAGATCTTTTCTTCATGATTTGTGAATCTATGTTAAAA 939
QY 291 MetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsnGlyVal 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 940 ATGGGAATATTAATGTTCTTACAGATATGAAGTGAATTAAGGAGAAATTTGATGATT 999
QY 311 ValAsn 312
    ||||| |||||
Db 1000 GTCAT 1005

RESULT 13
US-09-938-842A-1457
: Sequence 1457, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: FILE REFERENCE: SAME, AND METHODS OF USE
: CURRENT APPLICATION NUMBER: US/09/938,842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 1457
: LENGTH: 1062
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana

```

```

US-09-938-842A-1457
Alignment Scores:
Pred. No.: 4.37e-74 Length: 1062
Score: 692.50 Matches: 152
Percent Similarity: 58.33% Conservative: 37
Best Local Similarity: 46.91% Mismatches: 120
Query Match: 43.17% Indels: 15
DB: 9 Gaps: 4

US-10-047-825-4 (1-313) x US-09-938-842A-1457 (1-1062)
QY 5 ThrIleuMetGlnCysLeuValAlaValSerLeuLeuSerCysValAlaHisAlaGlnLeu 24
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37 ACAATCTTAATCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 96
QY 25 SerProThrPheThrAlaSerSerCysProAsnLeuGlnSerIleValArgAlaAlaMet 44
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 ACCCTCACTTCTAGATAGGTCATGCTTAATGCTCACTAATCATCATCGAAGAAACATT 156
QY 45 ThrGlnAlaValAlaSerGlnIleArgMetGlyAlaSerLeuAlaArgLeuPhePheHis 64
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 GTAATGACTTAAGAGTCGACCTCTGATTCGCTGCGACATCTTCCTTCACTTCCAC 216
QY 65 AspCysPheValGlnGlyCysAspGlySerIleLeuLeuAsp-----AlaGly 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 GACTGCTTGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 276
QY 81 GlyGlyIleuThrAlaGlyProAsnLeuAsnSerValArgGlyPheGlyValIleAspThr 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 277 ACAGAGAAAGATGCGCTTTGGAAGCAAAATTCGCGGATTTCCAGTGAATGATGAGA 336
QY 101 IleuValAsnValAlaAlaAlaAlaCysProGlyValValSerCysAlaAspIleLeuAla 120
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 337 ATGAAGAGTGGGGTGGAGAGGCAATGCCAGAAACCTTTCATCCAGATATCTCAACC 396
QY 121 LeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGlyProThrTyrSerValProLeuGly 140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 397 ATTGACCTCAACATCTGCTCACTTGGCAGAGAGTCTCTTCTTGAGAGGATCTTGGGA 456
QY 141 ArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProThr 160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 457 AGGAGAGCACTTTACAAAGCATTCCTGGAACCTCCCTAATGCAAAATCTCCAGCTCCATTC 516
QY 161 AlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnIleuSer---ProArgAsp 179
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 517 TTACACTTCCACAACCTTAAGCCAGCTTCAGAAATGATGATGATGATGATGATGATGAT 576
QY 180 MetThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArgGly 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 577 CTGCTGCTCTCCGGTGGTGCACACATTTGGTAAATCATGCTCATGTTATTCCTTGAC 636
QY 200 ArgIleTyrGly-----AspThrAspIleAsnAlaSerPheAla 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 637 AGATTATTAATTCACAGACACAGGTTTACCCGACCTTCACTCACTCACTACTTACTCTC 636
QY 213 AlaLeuArgGlnIleThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAsp 232
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 697 CAACCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 756
QY 233 ValGlnThrProValArgPheAspThrAlaTyrPheThrAsnLeuLeuSerArgGly 252
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 757 CTAGTACGCTCAGGTTTTCGACAAACAAATACTAGCTGAATTCNAAGAGCAAAAGGT 816
QY 253 LeuPheHisSerAspGlnIleuPhe-----AsnGlySerGlnAspAlaLeu 269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 817 CTATTCAGACAGCAGCAAGGTTGTTCTTACGCCCAATCCACTGACACAAATCCCTTGG 876
QY 270 ValArgGlnTyrSerAlaSerAlaSerLeuPheAsnAlaAspPheValAlaAlaMetIle 289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 877 GTGAGACATATGCTGATGCGACACAAACATTTCTTCAATGATTTTGGAGGCAATGAAT 936
QY 290 ArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsnGly 309

```

```
DB 937 AGATTGGGAACATTACCACTACAGCACTCAAGACAAATTCAGATTGAACGTAG 996
|||||
QY 310 ValValaInsSer 313
|||||
DB 997 GTTGTGAATCC 1008

RESULT 14
US-09-938-842A-2559
; Sequence 2559, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT1300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2559
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2559

Alignment Scores:
Pred. No.: 6,096-72 Length: 1011
Score: 674.50 Matches: 137
Percent Similarity: 62.21% Conservative: 54
Best Local Similarity: 44.61% Mismatches: 99
Query Match: 42.05% Indels: 17
DB: 9 Gaps: 5

US-10-047-825-4 (1-313) x US-09-938-842A-2559 (1-1011)
QY 24 LeuSerProThrPheTyrAlaSerSerProAsnLeuGlnSerIleValAlaAla 43
|||||
DB 88 TTAAACCTAGATTATTACAGCTACTGTCACACCGATTATTCAGCTCATAGAAAGAA 147
|||||
QY 44 MetTrpGlnAlaValAlaSerGlnIleArgMetGlyAlaSerLeuLeuArgLeuPhe 63
|||||
DB 148 ATGAATGCAATGATGAGAGAGATCTTGAATGCAAGCAATATTCGTCCTCACCTC 207
|||||
QY 64 HisAspCysPheValGlnGlyCysAspGlySerIleLeuLeuAsp-----Ala 79
|||||
DB 208 CACGACTGCTTTGCTCAAGATGATGATGATGCTGCTAGACGACAGACAAACTCTA 267
|||||
QY 80 GlyGlyGluValThrAlaGlyProAsnLeuAsnSerValArgGlyPheGluValIle 99
|||||
DB 268 CAGGAGAGAAAGAAAGCTTCTCCACACTAAATTCATTGAAGAGATACAAAATTGTGCAC 327
|||||
QY 100 ThrIleValArgAsnValGlnAlaAlaCysProGlyValValSerCysAlaAspIle 119
|||||
DB 328 AGAATCAAGAACATATTAATCCGAATGCTCTGAGATTGTTTACGCGCATGCTCTC 387
|||||
QY 120 AlaLeuAlaAlaValArgAspGlyThrAsnLeuGlnGlyGlyProThrTrpSerValPro 139
|||||
DB 388 ACAATTTGGTGTGAGATGCTACAAATCCTGGGTGGCCCTTACTGGAATGCTCTGTG 447
|||||
QY 140 GlyValArgAspSerThrThrAlaSerAlaSerIleValAlaAsnSerAspProProPro 159
|||||
DB 448 GGAAGAAAAGATTCACAAAACCGCAAGCTACGAGCTTGCCACACAAACACTTCACACTCA 507
|||||
QY 160 ThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArgAsp 179
```

```
DB 508 GAAGAGGTTTATACAGCATCATTCGTAGTCTATTCCTCAAGGCTCTCGGT-9AAGAC 567
|||||
QY 180 MetThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaValGlyCysThrThrPheArgGly 199
|||||
DB 568 ATGGTCGCTCTTATAGAGACGACAGATCGGAAAGACAAATGTTCGCAACTTCGATCC 627
|||||
QY 200 ArgIleTyrGlyAsp-----ThrAspIleAsnAlaSerPheAla 212
|||||
DB 628 CGAATTTATGAGATTTTTCACACTGACGTCACGCCCTTAATCCAGTTTCGAGACGACTTG 687
|||||
QY 213 AlaLeuArgGlnGlnThrCysProArgSerGlyGlyAspGly-----AsnLeuAlaPro 230
|||||
DB 688 GCAAGCTTCGAGAGATTTGTCGCGGAGTACGCGAAGAGATGATAGTAACTGACGCGC 747
|||||
QY 231 IleAspValGlnThrProValArgPheAspThrAlaTyrPheThrAsnLeuSerArg 250
|||||
DB 748 ATGACAAATGTGACGGCCGAACTCTTCATCTGATCTGACACACACTGTAAGAGA 807
|||||
QY 251 ArgGlyLeuPheHisSerAspGlnIleLeuPheAsp-----GlyGlySerGlnAsp 267
|||||
DB 808 GAAGGTTTACTGAAATTCGACAGAGATGTCACAGACTTCTTCGGATACAAACGGCG 867
|||||
QY 268 AlaLeuValArgGlnIleTyrSerAlaSerAlaSerLeuPheAsnAlaAspPheValAlaAla 287
|||||
DB 868 CGAATCGTGACAGATATCGCGAGATCCAGTGGCTTCTTCGAGCAATTCGAAAGTCG 927
|||||
QY 288 MetIleArgMetGlyAsnValGlyValLeuThrGlyThrAla---GlyGlnIleArgArg 306
|||||
DB 928 ATGTAAGATGTGGGAACATTTTGAACTGTGAAGACTTGCGTAGAGAGATTGAGAGA 987
|||||
QY 307 AsnCysArgValValaInsSer 313
|||||
DB 988 AATTCGAGATTTCGATAC 1008

RESULT 15
US-09-938-842A-1518
; Sequence 1518, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT1300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1518
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1518

Alignment Scores:
Pred. No.: 5,046-70 Length: 987
Score: 658.50 Matches: 145
Percent Similarity: 59.26% Conservative: 47
Best Local Similarity: 44.75% Mismatches: 109
Query Match: 41.05% Indels: 23
DB: 9 Gaps: 8

US-10-047-825-4 (1-313) x US-09-938-842A-1518 (1-987)
QY 9 CysLeuValAlaValSerLeuLeuSerCysVal-----AlaHisAlaGln----- 23
```

Db 22 TGGTTATTA---ATAACCTTCCTTAATGTCATCATCTCGGTTATGCGCAAGCCAC 78  
OY 24 -----LeuSerProThr-----PheTyrAlaSerCysProAsn 35  
Db 79 GCAAGGCGCTGCTCGTCTTCCGCACTAGAGATTGGGTTTACTTAAGTCTCTAGA 138  
OY 36 LeuGlnSerIleValArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGly 55  
Db 139 GCTGAGACTATTGTCGAACCGCGTGAATGCTGGTTTCAGTTCGACCCCAATCGCA 198  
OY 56 AlaSerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIle 75  
Db 199 CCCGGAATCTAGAAATGCAATTCACAGACTGTTTGTCCAAAGTGTGACGGTTGATC 258  
OY 76 LeuLeuAsp---AlaGlyGlyGluValThrAlaGlyProAsnLeuAsnSerValArgGly 94  
Db 259 CTTATATCCGAGCTAACACCGAGAGACCGCGGTCGCAACTCAAT---CTCCAAAGSA 315  
OY 95 PheGluValIleAspThrIleLysArgAsnValGluAlaAlaCysProGlyValAlaSer 114  
Db 316 TTTGAAGTATAGACAACGCCAAACGCAAGCTGAAGCCGCGTCTGAGTGTCTCT 375  
OY 115 CysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyProThr 134  
Db 376 TGTGCTGATATTTTACATTAAGCGGCTGACACAGTCATCTCTCACTCAAGAAACAGGC 435  
OY 135 TrpSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSer 154  
Db 436 TGGCAAGTACCAACGAGGAGTAGAGATGATAGATTCTTTGGCTTCGAATCTAAC--- 492  
OY 155 AsnProProProProThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGly 174  
Db 493 AATCTCCCTGGTCCCGTACTCCGTTGCCGTTCACAACAAGAAATTCCTCCGCTTGGA 552  
OY 175 LeuSerProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCys 194  
Db 553 CTCAATACGTCGATCTCTGCTCGTCTGTCGAGGACACAGATCGAAGACAGAGATGT 612  
OY 195 ThrThrPheArgGlyArgIleTyr-----GlyAspThrAspIleAsn 208  
Db 613 GGTGTATTCAGAAACAGGCTATTCAATACACCGCAACCGCAGATCCAACTCGAC 672  
OY 209 AlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyGlyAspGlyAsnLeu 228  
Db 673 CCAACATTTTGGCGAGCTTCAAACCAATGTCCCAAAACGGCAGGTTCACTGCCG 732  
OY 229 AlaProIleAspValGlnThrProValArgPheAspThrAlaTyrPheThrAsnLeu 248  
Db 733 GTGGATCTCGACACCGGAAGCGGAAGCACTTGGACACTTCTTACTACACAACTTAAGC 792  
OY 249 SerArgArgGlyLeuPheHisSerAspGlnGlnLeuPheAsnGlyGlySerGlnAspAla 268  
Db 793 CGTGGCGGTGAGTCTCCATCGATCGATCAAGCTTTGGACCGATCCAGCCACTAGACC 852  
OY 269 LeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsnAlaAspPheValAlaAlaMet 288  
Db 853 ATAGTGCACAACTTGATGGCTCTCTAGAAAGCACCTTCAACGTTGAGTTGCGAGGCTAATG 912  
OY 289 IleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArgAsnCys 308  
Db 913 GTCAGAAATGATATATGCTGTGTACCGAGCTAATGGGAAATCCGTAGGGTTTGC 972  
OY 309 ArgValValAsn 312  
Db 973 TCTGCGGTTAAT 984

Search completed: May 3, 2003, 13:53:55  
Job time : 114 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 3, 2003, 13:44:59 ; Search time 1083 Seconds

(without alignments)  
4680.690 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 313

Sequence: 1 MASPTMQCLVAIVSLSCVA.....VGVLTGATGAGTIRRCRVNS 313

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Word size: 16

Total number of hits satisfying chosen parameters: 376

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:  
-MODE=frame+pn\_model -DEV=xlh  
-O/cp2n2.1/USPNO.spool/US10047825/runat\_26042003.112042.7654/app.query.fasta.1.455  
-DB=EST -QMT=fastap -SUFIT=first -MINMATCH=0.1 -DOOPCL=0 -DOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=quality -THR.MIN=16 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc  
-NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10047825 @CGN.1.1.1456 @runat.26042003.112042.7654 -NCPU=6 -ICPU=3  
-NO\_XIPYX -NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlmu:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estcl:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_yrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_tod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	34.8	538	12	BG842311
2	94	30.0	848	9	AA979912
3	77	24.6	446	12	BG841858
4	46	14.7	850	9	A1374530
5	37	11.8	230	14	T18410
6	37	11.8	369	10	BE366146
7	37	11.8	380	12	BG411665
8	37	11.8	434	10	AW287460
9	37	11.8	476	13	BM324451
10	37	11.8	479	10	BE601026
11	37	11.8	480	10	BE366914
12	37	11.8	480	10	BE600167
13	37	11.8	487	10	BE599526
14	37	11.8	495	10	BE600529
15	37	11.8	522	10	BE594365
16	37	11.8	524	10	BE597461
17	37	11.8	533	10	BE125306
18	37	11.8	537	10	BE595535
19	37	11.8	555	10	AW671012
20	37	11.8	560	10	BE600110
21	37	11.8	561	10	BE361738
22	37	11.8	564	10	BE600142
23	37	11.8	566	10	BE366224
24	37	11.8	566	10	BE600823
25	37	11.8	569	10	BE357554
26	37	11.8	570	10	BE358471
27	37	11.8	589	10	BE597430
28	37	11.8	591	10	BE362356
29	37	11.8	593	10	BE362200
30	37	11.8	615	10	BE361528
31	37	11.8	633	10	AW671673
32	37	11.8	591	12	BG462797
33	37	10.2	602	12	BG463106
34	37	10.2	625	10	AW672124
35	37	10.2	625	10	BE125808
36	29	9.3	452	12	BG649323
37	26	8.3	501	13	BM318238
38	25	8.0	404	10	BE362420
39	25	8.0	436	10	BE597204
40	25	8.0	471	10	BE362317
41	25	8.0	471	10	BE597234
42	25	8.0	483	13	BM318212
43	25	8.0	512	10	AW672196
44	25	8.0	566	10	BE600412
45	25	8.0	570	10	BE597297

#### ALIGNMENTS

RESULT 1  
BG842311  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

538 bp mRNA linear EST 29-MAY-2001  
MEST29-D08.T3 ISUM4-TN zea mays cDNA clone MEST29-D08 3', mRNA  
sequence.  
BG842311  
BG842311.2 GI:14244343  
EST.  
Zea mays.  
Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
Clade; Panicoideae; Andropogoneae; Zea.  
Qiu,F., Cui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.



TITLE  
JOURNAL  
COMMENT

Expressed Sequence Tags from B73 Maize Seedlings and Silks  
Unpublished (2001)  
On May 25, 2001 this sequence version replaced gi:14208633.  
Contact: Patrick S. Schnable  
Schnable Laboratory  
Iowa State University  
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA  
Tel: 515-294-0975  
Fax: 515-294-2299  
Email: schnable@iastate.edu

## FEATURES

source  
1. 538  
/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="MEST29-D08"  
/clone\_1lb="ISUM4-TN"  
/tissue\_type="Seedling and silk"  
/lab\_host="DH10B"  
/note="Vector: pRT73PAC; Site\_1: EcoRI; Site\_2: NotI;  
ds-cDNA molecules were generated as follows. First-strand  
cDNA was prepared from oligo-dT selected mRNA by priming  
with a NotI oligo-dT primer (5'  
AACTGGAGATTCGGCGCGGCGGAGATTTTCTTTTCTTTT). The  
resulting DNA:RNA hybrid was treated with RNase H and used  
as a template for DNA Pol-catalyzed second strand  
synthesis. After the addition of EcoRI adaptors, the  
ds-cDNAs were digested with NotI and size-selected. The  
resulting molecules were directionally cloned into the  
EcoRI and NotI sites of the pRT73PAC vector. The library  
then went through one round of normalization to COT value  
of 5 based on the methods of Marcelo Bento Soares (Genome  
Research 6: 791-806, 1996)."  
BASE COUNT 121 a 162 c 153 g 102 t  
ORIGIN

## Alignment Scores:

Pred. No.: 5,66e-95 Length: 538  
Score: 109.00 Matches: 109  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 34.82% Indels: 0  
DB: 12 Gaps: 0

US-10-047-825-4 (1-313) x BG842311 (1-538)

QY 205 ThrAsp1LeaSnAlaSerPheAlaLeuAArgGInGInThrCysProArgSerGlyGly 224  
|||||  
DB 3 ACCGACATCAAGCCTCTTCGCGGCGCTCGCGGACGACAGCTGCCGCGGCGCGC 62  
QY 225 AspglyAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyPhe 244  
|||||  
DB 63 GACGGCAACCTGGCGCCCTTCGACGTCGACGCGCGGTGAGGTTGACACGCGCTACTTC 122  
QY 245 ThrAsnLeuSerArgArgGlyLeuPheHisSerAspGlnGlnLeuPheAsnGlyGly 264  
|||||  
DB 123 ACCAACCTCTCTGCGGCGGCGCTGCTTCCTCGACGACGACGCTCTTCAACGCGCGG 182  
QY 265 SerGlnAspAlaLeuValArgGlnTySerAlaSerAlaSerLeuPheAsnAlaAspHe 284  
|||||  
DB 183 TCGCAGAGAGCGCTGGTGAGGACATACAGCGCCGCTCTTCACCGCGACTTC 242  
QY 285 ValAlaAlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIle 304  
|||||  
DB 243 GTGGCGACCATGATTAGATGGGCAACGTTGGGTCGTACCGGACCGCTGACAGATC 302  
QY 305 ArgArgAsnGlyArgValValAsnSer 313  
|||||  
DB 303 AGCGCACTGCGCGGCGCTGCAACAGC 329

RESULT 2  
AA979912 848 bp mRNA linear EST 26-MAY-1998  
LOCUS  
DEFINITION  
MEST3-D6.tw1412.Seq ISUM2 Zea mays cDNA clone MEST3-D6 5', mRNA  
sequence.

ACCESSION  
AA979912  
VERSION  
AA979912.1 GI:3157290  
KEYWORDS  
EST.  
SOURCE  
Zea mays.  
ORGANISM

REFERENCE  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 848)  
Wen,T.J., Ashlock,D.A. and Schnable,P.S.  
Expressed Sequence Tags from B73 Maize Seedlings  
Unpublished (1997)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

contact: Schnable, PS  
Schnable Laboratory  
Iowa State University  
G405 Agronomy, Ames, IA 50011, USA  
Tel: (515)-294-0975  
Fax: (515)-294-2299  
Email: schnable@iastate.edu  
PCR Primers  
FORWARD: tw1412 (5'-GAAGATACCCGACCAACC-3')  
BACKWARD: 77-YI (5'-TAATCGACTACTATTAGGC-3')  
Plate: MEST3 row: D column: 6  
Seq primer: tw1412 (5'-GAAGATACCCGACCAACC-3').  
location/Qualifiers  
1. 848

## FEATURES

source  
1. 848  
/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="MEST3-D6"  
/clone\_1lb="ISUM2"  
/tissue\_type="above ground tissues"  
/dev\_stage="Two-leaf-stage green seedling"  
/note="Organ: green seedlings; Vector: pAD-GAL4; Site\_1:  
EcoRI; Site\_2: XhoI; ds-cDNA molecules were generated as  
follows. First-strand cDNA was prepared from oligo-dT  
selected mRNA by priming with an XhoI oligo-dT primer. The  
resulting DNA:RNA hybrid was treated with RNase H and used  
as a template for DNA Pol-catalyzed second strand  
synthesis. After the addition of EcoRI adaptors, the  
ds-cDNAs were digested with XhoI and size-selected. The  
resulting molecules were directionally cloned into the  
EcoRI and XhoI sites of the HybridZAP lambda vector  
(Stratagene) and excised as pAD-GAL4 phagemids."  
BASE COUNT 162 a 284 c 236 g 147 t 19 others  
ORIGIN

## Alignment Scores:

Pred. No.: 2.89e-80 Length: 848  
Score: 94.00 Matches: 163  
Percent Similarity: 98.79% Conservative: 0  
Best Local Similarity: 98.79% Mismatches: 1  
Query Match: 30.03% Indels: 2  
DB: 9 Gaps: 0

US-10-047-825-4 (1-313) x AA979912 (1-848)

QY 87 ProAsnLeuSnsrValArgGlyPheGluValIleAspThrIleLysArgAsnValGlu 106  
|||||  
DB 1 CCGAAGCTGAAGCTGGCGCGGCTTGGAGCTATCACCACCAAGCGAAGCTGAG 60  
QY 107 AlAlaIacysProGlyValValIalserCysAlaAspIleLeuAlaLeuAlaAlaArgAspGly 126  
|||||  
DB 61 GCCCGTGGCCCGCGCGTGTGTCGCGGACATCTCGCGCTTCCGCGCGCGACGGA 120  
QY 127 ThrAsnLeuGlnGlyProThrTrpSerValProLeuGlyValArgAspSerThrThr 146



```

Db 121 ACCAAGCTTCGCGGCGGAGAGCTGAGAGTCCGCTCGGGGGGAGGACATCGAGCAG 180
QY 147 A1aSerAlaSerleuAlaasnSerAsnPro-PropioProthAlaSerleuGlyThrle 166
Db 181 GCCAGGCGCCGCTCGCCACACACCA-CTTCCCGCCCGGACGCGCAGCTCGGACGCT 239
QY 166 u1eSerleuPheGlyArgGlnGlyLeuSerProArgAspMetThrAlaLeuSerGlyAl 186
Db 240 CATCTCCCTTTGCGAGGAGGAGGCTGTGCGCGCGACATGACGCGCTGTGCGGCGCC 239
QY 186 aH1ThrIleGlyAlaAlaArgCysThrThrPheArgGlyArgIleGlyAspThrAs 206
Db 300 GCACACCATCGGCGAGCGCGGTGACACACCTTCGCGCGCGCATCTACGCGCAGCAGCA 359
QY 206 p1leAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyGlyAsp 226
Db 360 CATCAAGCCTCTCGCGCGCGGTGCGGCGACAGACGTGCGCGGTGCGGCGCGCAGG 419
QY 226 yAsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyrPheThrAs 246
Db 420 CAACCTGCGCGCCATCGACGTGACAGCGCGGTGAGTTGACAGCGCGCTACTTACCA 479
QY 246 nLeuLeuSerArg 250
Db 480 CCTGCTGTCCGCG 492

```

## RESULT 3

```

Bg841858/c 446 bp mRNA linear EST 29-MAY-2001
LOCUS MEST27-D09.T3 ISUM4-TN Zea mays cDNA clone MEST27-D09.3', mRNA
DEFINITION
Sequence.
ACCESSION Bg841858
VERSION Bg841858.2 GI:14244122
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 446)
Qiu,F., Cui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14208180.
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu
PCR Primers
FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
Location/Qualifiers
1..446

```

## FEATURES

SOURCE

```

/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST27-D09"
/clone_1b="ISUM4-TN"
/tissue_type="Seedling and silk"
/lab_host="DH10B"
/note="Vector: pT733PAC; Site.1: EcoRI, Site.2: NotI;
ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'
AAGTGAAGAAATTCGCGCGCGGTGACACACCTTCGCGCGCATCTACGCGCAGCAGCA
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the

```

ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT733PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

BASE COUNT 65 a 147 c 166 g 68 t

## ORIGIN

## Alignment Scores:

Pred. No.:	Length:	Matches:
3,94e-64	446	139
Score:	77.00	Conservative: 0
Percent Similarity:	97.20%	Mismatches: 2
Best Local Similarity:	97.20%	Indels: 4
Query Match:	24.60%	Gaps: 0
DB:	12	

US-10-047-825-4 (1-313) x Bg841858 (1-446)

```

QY 53 ArgMetGlyAlaSerleuLeuArgLeuPhePheHisAspCysPheValGlnGlyCysAsp 72
Db 422 AGAGTGGGGGCTCTGCTGCTGAGGCTTCTTCCAGCACTGCTTCAAGGCTGCGAC 364
QY 73 GlySerIleLeuLeuAspAlaGlyGlyGlyThrAlaGlyProAsnLeuAsnSerVal 92,
Db 363 GGATGATCTCTTCGACGCGGAGGAGAGAGACCGCGGCGGCGGACCTAAGCTGAGTG 304
QY 93 ArgGlyPheGlnValIleAspThrIleYsArgAsnValGlnAlaAlaCysProGlyVal 112
Db 303 CCGCGCTTTGAGGTGATGACACCATCAAGCGAGACGTAGAGCGCGCTGCCGCGCTC 244
QY 113 ValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGlyThrAsn-LeuLeuGlyG 132
Db 243 GGTGCTGGCGCGACATCCTCGCGCTTCCGCGCGCGCGGCGGCGGCGGCGGCGGCGG 185
QY 132 yProthrrPpSerValProLeuGlyArgArgAspSerThrThrAlaSerAlaSerleuAl 152
Db 184 GCCGAGCTTGAGAGCTGACGCGCTGCGGCGGCGGAGCTGACGAGCGGCGGCTGCTGC 125
QY 152 aAsnSerAsnPro-PropioProthrrAlaSerleuGlyThrLeuLeuSerleuPheGly 172
Db 124 CACACGCA-CTTCCCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 66
QY 172 rgGlnGlyLeuSerProArgAspMetThrAlaLeuSerGlyAlaHisThrIleGlyGln 192
Db 65 GCGAGGCGCTGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6
QY 192 lAArg 193
Db 5 CCGCG 1

```

## RESULT 4

```

AI374530/c 850 bp mRNA linear EST 21-JAN-1999
LOCUS MEST3-D6.POLYT-N-Seq ISUM2 Zea mays cDNA clone MEST3-D6.5', mRNA
DEFINITION
Sequence.
ACCESSION AI374530
VERSION AI374530.1 GI:4174550
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 850)
Wen,T.J., Ashlock,D.A. and Schnable,P.S.
Expressed Sequence Tags from B73 Maize Seedlings
Unpublished (1997)
Contact: Schnable, PS
Schnable Laboratory
Iowa State University
G405 Agronomy, Ames, IA 50011, USA
Tel: (515)-294-0975

```

Fax: (515)-294-2299  
Email: schmaleel@state.edu

PCR Primers  
FORWARD: tw1412 (5'-GAAGTACCCACCAAGC-3')  
BACKWARD: tw1412 (5'-TAATGACACTCATAGGC-3')

Plate: MEST3 Row: D Column: 6  
Seq primer: POLYT-N (5'-TTTTTTTTTTTTTTTTTTTTTTT (AGC)-3').

Location/Qualifiers

## FEATURES

source

1. 850  
/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone\_1lb="150M2"  
/tissue\_type="above ground tissues"  
/dev\_stage="Two-leaf-stage green seedling"  
/lab\_host="XLI-MFR Blue"  
/note="Organ: green seedlings; Vector: PAD-GAL4; Site\_1: EcoRI; Site\_2: XhoI; ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with an XhoI oligo-dT primer. The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with XhoI and size-selected. The resulting molecules were directionally cloned into the EcoRI and XhoI sites of the HybridZAP Lambda vector (Stratagene) and excised as PAD-GAL4 phagemids."

BASE COUNT 168 a 226 c 210 g 194 t 52 others

ORIGIN

## Alignment Scores:

Pred. No.: 7.17e-34 Length: 850  
Score: 46.00 Matches: 46  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 14.70% Indels: 0  
DB: 9 Gaps: 0

US-10-047-825-4 (1-313) x A1374530 (1-850)

QY 268 AAlaValArgSerLeuSerLeuPheAsnAlaSerValAlaAla 287  
|||||  
DB 407 GCGCTGTGAGGACGACGCGCGCTCTTCAAGCCGACCTTCGTGGCAGCC 348  
QY 288 MetIleArgMetGlyAsnValGlyValLeuPheGlyThrAlaGlyGlnIleArgArgAsn 307  
|||||  
DB 347 ATATTAGGATGGCAACGTTGGGTGCTACCGCAGCTGACAGATCAGCGCAGC 288  
QY 308 CysArgValAlaAsnSer 313  
|||||  
DB 287 TGCCGGGTCTGTCAACAGC 270

## RESULT 5

T18410

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST

Ze mays

Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 230)

Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T., Helentjaris,T., Baysdorfer,C., Almita,E., Ferl,R., Hadden,U. and

Latinis,B.

Partial sequencing and mapping of clones from two maize cDNA libraries

Plant Mol. Biol. 26, 1085-1101 (1994)

Journal

## MEDLINE

9511093

## COMMENT

Contact: The Maize cDNA Project

Helentjaris TG (primary contact)

Dept. of Plant Sciences

University of Arizona

Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721

ph: 602-6218-746

ph: 602-621-7186

E-mail: helentjaris@ccit.arizona.edu

Chris Baysdorfer

Department of Biological Sciences, School of Science

California State University, Hayward

Hayward, CA 94542

ph: 510-881-3459

fax: 510-727-2035

E-mail: cbaysdorfer@csuhayward.edu

Rob Ferl

Interdisciplinary Center for Biotechnology Research

DNA Sequencing Core

University of Florida

P.O. Box 100695

Gainesville, FL 32611-0695

ph: 904-392-1928, ext. 301

fax: 904-392-4072

E-mail: robferl@nervm.ncrc.ufl.edu

## FEATURES

source

Location/Qualifiers

1. 230

/organism="Zea mays"

/strain="B73"

/db\_xref="taxon:4577"

/clone\_1lb="6C02d10"

/clone\_1lb="etiolated seedling"

/lab\_host="DH10B"

/note="Vector: Z1plox; Site\_1: SalI; Site\_2: NotI; oligo-dT selected mRNA was prepared from above-ground tissue from 10 day-old etiolated seedlings. ds-cDNA was prepared by priming with a NotI oligo-dT oligomer and the second strand by RNase-nicking of the DNA:RNA hybrid with DNA PolI fill-in. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the Z1plox phage vector, excised as plasmids, and then individually analyzed."

BASE COUNT 37 a 82 c 60 g 48 t 3 others

## ORIGIN

## Alignment Scores:

Pred. No.: 8.95e-26 Length: 230  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.82% Indels: 0  
DB: 14 Gaps: 0

US-10-047-825-4 (1-313) x T18410 (1-230)

QY 1 MetAlaSerProThrLeuMetGlnCysLeuValAlaValSerLeuSerCysValAla 20  
|||||

DB 58 ATGGCGTCTCCACCTTGATGCAATGCTGCTCCCTTCCTCTCTGTCGCC 117  
|||||

QY 21 HisAlaGlnLeuSerProThrPheTyrAlaSerSerCysProAsnLeuGln 37  
|||||

DB 118 CAGCCACAGCTCTGCGCCACGCTTATGCTCTCTCCGCCCAACCTGCAG 168  
|||||

## RESULT 6

BE36146

LOCUS

DEFINITION

BE36146

369 bp mRNA linear EST 20-JUL-2000

PI1\_31\_A02\_b1\_A002 pathogen induced 1 (PI1) Sorghum bicolor cDNA,

mRNA sequence.

ACCESSION BE366146  
 VERSION BE366146.1 GI:9307703  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 369)  
 Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt L.H.  
 An EST database from Sorghum: pathogen-induced plants  
 TITLE An EST database from Sorghum: pathogen-induced plants  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@prattuga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16..The threshold for highest quality sequence is 20.  
 Seq primer: JEN REV  
 High quality sequence start: 180  
 High quality sequence stop: 362  
 POLYA-No.

# FEATURES

source Location/Qualifiers  
 1..369  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Patxon:4558"  
 /note="Organ: Anthracnose-infected leaves from two-week-old sorghum plants 48 hr after inoculation; Vector: Bluescript II from Lambda Zap II; Site.1: XhoI; Site.2: EcoRI; Two-week-old sorghum plants (BRX 623 cultivar) were infected with pathogen (isolate PM421 of Colletotrichum graminicola, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptibility to anthracnose disease. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."  
 BASE COUNT 67 a 125 c 98 g 79 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 1.48e-25 Length: 369  
 Score: 37.00 Matches: 37  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.82% Indels: 0  
 DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE366146 (1-369)

OY 52 GlnarGmetGlyAlaSerLeuLeuAArgLeuPhePheHisAspCysPheValGlnGlyCys 71  
 |||||||  
 Db 210 CAGAGATGGGCGGCTCTCTGCTCAGGCTCTCTCCACAGACTGTTGTCAAGGTTC 269  
 |||||||

OY 72 AspGlySerIleLeuLeuAspAlaGlyGlyGluLysThrAlaGlyProAsn 88  
 |||||||  
 Db 270 GACGGCTCATCTTCGACGCGGAGGAGAGACGCGGCCCAAC 320  
 |||||||

RESULT 7  
 BG411665  
 LOCUS BG411665 380 bp mRNA linear EST 13-MAR-2001

DEFINITION EMI\_60\_C03.bl\_A002 Embryo 1 (EMI) Sorghum bicolor cDNA, mRNA sequence.  
 ACCESSION BG411665  
 VERSION BG411665.1 GI:13317218  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 380)  
 Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.  
 An EST database from Sorghum: developing embryos  
 TITLE Unpublished (2000)  
 JOURNAL Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@prattuga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20.  
 Seq primer: JEN REV  
 High quality sequence start: 342  
 High quality sequence stop: 342  
 POLYA-No.

# FEATURES

source Location/Qualifiers  
 1..380  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Embryo 1 (EMI)"  
 /note="Organ: Embryos germinated for 24 hr; Vector: Bluescript II from Lambda Zap II; Site.1: XhoI; Site.2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."  
 BASE COUNT 67 a 128 c 107 g 78 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.53e-25 Length: 380  
 Score: 37.00 Matches: 37  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.82% Indels: 0  
 DB: 12 Gaps: 0

US-10-047-825-4 (1-313) x BG411665 (1-380)

OY 52 GlnarGmetGlyAlaSerLeuLeuAArgLeuPhePheHisAspCysPheValGlnGlyCys 71  
 |||||||  
 Db 198 CAGAGATGGGCGGCTCTCTGCTCAGGCTCTCTCCACAGACTGTTGTCAAGGTTC 257  
 |||||||

OY 72 AspGlySerIleLeuLeuAspAlaGlyGlyGluLysThrAlaGlyProAsn 88  
 |||||||  
 Db 258 GACGGCTCATCTTCGACGCGGAGGAGAGACGCGGCCCAAC 308  
 |||||||

RESULT 8  
 AW287460  
 LOCUS AW287460 434 bp mRNA linear EST 19-JUL-2000  
 DEFINITION LG1\_228.C02.bl\_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA sequence.  
 ACCESSION AW287460  
 VERSION AW287460.2 GI:6859454  
 KEYWORDS EST.

ORGANISM sorghum.  
 SOURCE Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 434)  
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.

**TITLE** An EST database from Sorghum: light-grown seedlings  
**JOURNAL** Unpublished (2000)  
**COMMENT** On Jan 6, 2000 this sequence version replaced gi:6677304.  
 Contact: Cordomier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
 Seq primer: JEN REV  
 High quality sequence stop: 405  
 POLYA-No.

**FEATURES**  
**source** Location/Qualifiers

1..434  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_1lb="Light Grown 1 (LG1)"  
 /note="Organ: 10- to 14-day-old light-grown (greenhouse) seedlings; Vector: Lambda Zap; Site\_1: XhoI; Site\_2: EcoRI ; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."  
 BASE COUNT 71 a 152 c 125 g 86 t  
 ORIGIN

**Alignment Scores:**

Pred. No.: 1.76e-25 Length: 434  
 Score: 37.00 Matches: 37  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.82% Indels: 0  
 DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x AM287460 (1-434)

QY 52 GlnArgMetGlyAlaSerLeuLeuArgLeuPheHisAspCysPheValGlnGlyCys 71  
 |||||||  
 DB 186 CAGAGATGGGCGGCTCTGCTCAGGCTCTTCACAGATGCTTTTAAAGTTGC 245

QY 72 AspGlySerIleLeuLeuAspAlaGlyGlyGluLysThrAlaGlyProasn 88  
 |||||||  
 DB 246 GACGGCTGATCCTTCGACGCGGAGAGAGACGCGGCCCAAC 296

**RESULT 9**

BM324451 476 bp mRNA linear EST 04-JAN-2002  
 LOCUS PIC1\_31.D06.b1\_A002 pathogen-infected compatible 1 (PIC1) Sorghum  
 bicolor cDNA, mRNA sequence.  
 ACCESSION BM324451  
 VERSION BM324451.1 GI:18063120  
 KEYWORDS EST.

**SOURCE**

ORGANISM sorghum.  
 Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 476)  
 Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wang,R.,  
 Sudman,M. and Pratt,L.H.  
 An EST database from Sorghum: plants infected with a compatible  
 pathogen

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

Unpublished (2002)  
 Contact: Cordomier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector, and regions below Phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with polyTmX or T7 sequencing primer, are presented as the reverse complement.  
 Seq primer: JEN REV  
 High quality sequence stop: 473  
 POLYA-No.

**FEATURES**  
**source** Location/Qualifiers

1..476  
 /organism="Sorghum bicolor"  
 /cultivar="Btx623"  
 /db\_xref="taxon:4558"  
 /clone\_1lb="Pathogen-Infected compatible 1 (PIC1)"  
 /tissue\_type="Leaves"  
 /dev\_stage="4-week-old seedlings infected with  
 Colletotrichum graminicola"  
 /note="Vector: pBluescript II SK(-) from Lambda Zap II;  
 Site\_1: XhoI; Site\_2: EcoRI; Four-week-old sorghum  
 seedlings were sprayed with spore suspension prepared from  
 3-week-old PM421, a sorghum isolate of the anthracnose  
 pathogen Colletotrichum graminicola. Inoculated plants  
 were kept in a 25 C dark growth chamber with 100% relative  
 humidity for 24 hr, followed by 12/12 hr of light/dark  
 cycle at 25 C with 90% relative humidity for another 24  
 hr. All leaves were harvested and quick frozen with liquid  
 nitrogen and stored in a -80 C freezer. The library was  
 made from poly-A RNA in the cloning vector lambda Zap II.  
 Clones to be sequenced were prepared by mass excision.  
 WARNING: While most or all ESTs are expected to derive  
 from the host plant, no effort was made to eliminate ESTs  
 deriving from the pathogen."  
 BASE COUNT 78 a 164 c 138 g 96 t  
 ORIGIN

**Alignment Scores:**

Pred. No.: 1.94e-25 Length: 476  
 Score: 37.00 Matches: 37  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.82% Indels: 0  
 DB: 13 Gaps: 0

US-10-047-825-4 (1-313) x BM324451 (1-476)

QY 52 GlnArgMetGlyAlaSerLeuLeuArgLeuPheHisAspCysPheValGlnGlyCys 71  
 |||||||  
 DB 205 CAGAGATGGGCGGCTCTGCTCAGGCTCTTCACAGATGCTTTTAAAGTTGC 264

QY 72 AspGlySerIleLeuLeuAspAlaGlyGlyGluLysThrAlaGlyProasn 88  
 |||||||  
 DB 265 GACGGCTGATCCTTCGACGCGGAGAGAGACGCGGCCCAAC 315

**RESULT 10**

BE601026 479 bp mRNA linear EST 18-AUG-2000  
 LOCUS P11\_96\_A03.b1\_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,  
 mRNA sequence.  
 ACCESSION BE601026  
 VERSION BE601026.1 GI:9856099  
 KEYWORDS EST.

**SOURCE**

ORGANISM sorghum.  
 Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 479)  
 Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt  
 ,L.H.  
 An EST database from Sorghum: pathogen-induced plants  
 Unpublished (2000)  
 Contact: Cordomier-Pratt MM

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

Unpublished (2000)  
 Contact: Cordomier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV

High quality sequence stop: 478  
POLYA-No.

## FEATURES

## SOURCE

Location/Qualifiers  
1..479

/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"

/clone\_lib="Pathogen induced 1 (PII)"  
/note="Organ: Anthracnose-infected leaves from two-week-old sorghum plants 48 hr after inoculation; Vector: pBluescript II from lambda Zap II; Site 1: XhoI; Site 2: EcoRI; Two-week-old sorghum plants (B7X 623 cultivar) were infected with pathogen (isolate PM421 of Colletotrichum graminicola, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptibility to anthracnose disease. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT 79 a 165 c 138 g 96 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1,95e-25 Length: 479  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.82% Indels: 0  
DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE601026 (1-479)

QY 52 GlnArgMetGlyAlaSerLeuArgLeuPhePheHisAspCysPheValGlnGlyCys 71  
|||||

Db 208 CAGAGATGGGCGCTCTGCTCAGGCTCTTCTCCACGACGCTTTGTTCAGAGTTGC 267  
|||||

QY 72 AspGlySerIleLeuLeuAspAlaGlyGlyIuLysThrAlaGlyProAsn 88  
|||||

Db 268 GACGGCTGCATCCTTCTGCAGCGGAGGAGAGACGCGCCGCCCAAC 318  
|||||

## RESULT 11

BE366914

LOCUS BE366914 480 bp mRNA linear EST 20-JUL-2000

DEFINITION P11.42.B10.b1.A002 Pathogen induced 1 (PII) Sorghum bicolor cDNA,

mRNA sequence.

ACCESSION BE366914

VERSION BE366914.1 GI:9308471

KEYWORDS EST.

SOURCE sorghum.

ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (Bases 1 to 480)

AUTHORS Cordomier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt

L.H.

An EST database from Sorghum: pathogen-induced plants

Unpublished (2000)

CONTACT: Cordomier-Pratt MM

LABORATORY for Genomics and Bioinformatics

## FEATURES

## SOURCE

Location/Qualifiers  
1..480

/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"

/clone\_lib="Pathogen induced 1 (PII)"  
/note="Organ: Anthracnose-infected leaves from two-week-old sorghum plants 48 hr after inoculation; Vector: pBluescript II from lambda Zap II; Site 1: XhoI; Site 2: EcoRI; Two-week-old sorghum plants (B7X 623 cultivar) were infected with pathogen (isolate PM421 of Colletotrichum graminicola, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptibility to anthracnose disease. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT 79 a 165 c 140 g 96 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1,95e-25 Length: 480  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.82% Indels: 0  
DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE366914 (1-480)

QY 52 GlnArgMetGlyAlaSerLeuArgLeuPhePheHisAspCysPheValGlnGlyCys 71  
|||||

Db 205 CAGAGATGGGCGCTCTGCTCAGGCTCTTCTCCACGACTGCTTGTTCAGAGTTGC 264  
|||||

QY 72 AspGlySerIleLeuLeuAspAlaGlyGlyIuLysThrAlaGlyProAsn 88  
|||||

Db 265 GACGGCTGCATCCTTCTGCAGCGGAGGAGAGACGCGCCGCCCAAC 315  
|||||

## RESULT 12

BE600167

LOCUS BE600167 480 bp mRNA linear EST 18-AUG-2000

DEFINITION P11.80.B04.b1.A002 Pathogen induced 1 (PII) Sorghum bicolor cDNA,

mRNA sequence.

ACCESSION BE600167

VERSION BE600167.1 GI:9855240

KEYWORDS EST.

SOURCE sorghum.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (Bases 1 to 480)

AUTHORS Cordomier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt

L.H.

An EST database from Sorghum: pathogen-induced plants

Unpublished (2000)

CONTACT: Cordomier-Pratt MM

Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@pratt.uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
Seq primer: JEN REV  
High quality sequence stop: 378  
POLYA-No.

## FEATURES

source

Location/Qualifiers

1..480

/organism="Sorghum bicolor"

/db\_xref="taxon:4558"

/clone\_lib="Pathogen induced 1 (PI1)"

/note="Organ: Anthracnose-infected leaves from

two-week-old sorghum plants 48 hr after inoculation;

Vector: pBluescript II from lambda Zap II; Site\_1: XhoI;

Site\_2: EcoRI; Two-week-old sorghum plants (BRX 623

cultivar) were infected with pathogen (isolate FRM421 of

Colletotrichum graminicola, which is a sorghum isolate).

RNA was prepared from infected leaves harvested from 45

seedlings (2 weeks old) exhibit juvenile resistant

reaction, which is an incompatible interaction. As they

grow older (4 weeks or older), plants resume susceptibility

to anthracnose disease. The library was made from poly-A

RNA in the cloning vector lambda Zap II. Clones to be

sequenced were prepared by mass excision. WARNING: While

most or all ESTs are expected to derive from the host

plant, no effort was made to eliminate ESTs deriving from

the pathogen."

BASE COUNT

84 a 161 c 138 g 97 t

ORIGIN

## Alignment Scores:

Pred. No.:	Length:	480
Score:	Matches:	37
Percent Similarity:	Conservative:	0
Best Local Similarity:	Mismatches:	0
Query Match:	Indels:	0
DB:	Gaps:	0

US-10-047-825-4 (1-313) x BE600167 (1-480)

QY 52 GlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCys 71

Db 217 CAGAGGATGGGGCCCTCTGCTCAGGCTCTTCACAGACTGCTTGTTCAGAGTTGC 276

QY 72 AspGlySerIleLeuLeuAspAlaGlyGlyLysThrAlaGlyProasn 88

Db 277 GACGGCTCGATCTTCTTCAGCCCGAGGGGAGAGACGGCCGCCAC 327

RESULT 13

BE599526 487 bp mRNA linear EST 18-AUG-2000

LOCUS P11.88.F01.b1.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,

DEFINITION mRNA sequence.

ACCESSION BE599526

VERSION BE599526.1 GI:9854599

KEYWORDS EST.

SOURCE

ORGANISM

Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoidae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 487)

AUTHORS Cordomier-Pratt, M.-M., Gingle, A., Dean, R., Sudman, M. and Pratt

TITLE An EST database from Sorghum: pathogen-induced plants

JOURNAL Unpublished (2000)

## COMMENT

Contact: Cordomier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@pratt.uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
Seq primer: JEN REV  
High quality sequence stop: 477  
POLYA-No.

## FEATURES

source

Location/Qualifiers

1..487

/organism="Sorghum bicolor"

/db\_xref="taxon:4558"

/clone\_lib="Pathogen induced 1 (PI1)"

/note="Organ: Anthracnose-infected leaves from

two-week-old sorghum plants 48 hr after inoculation;

Vector: pBluescript II from lambda Zap II; Site\_1: XhoI;

Site\_2: EcoRI; Two-week-old sorghum plants (BRX 623

cultivar) were infected with pathogen (isolate FRM421 of

Colletotrichum graminicola, which is a sorghum isolate).

RNA was prepared from infected leaves harvested from 45

seedlings (2 weeks old) exhibit juvenile resistant

reaction, which is an incompatible interaction. As they

grow older (4 weeks or older), plants resume susceptibility

to anthracnose disease. The library was made from poly-A

RNA in the cloning vector lambda Zap II. Clones to be

sequenced were prepared by mass excision. WARNING: While

most or all ESTs are expected to derive from the host

plant, no effort was made to eliminate ESTs deriving from

the pathogen."

BASE COUNT

81 a 167 c 143 g 96 t

ORIGIN

## Alignment Scores:

Pred. No.:	Length:	487
Score:	Matches:	37
Percent Similarity:	Conservative:	0
Best Local Similarity:	Mismatches:	0
Query Match:	Indels:	0
DB:	Gaps:	0

US-10-047-825-4 (1-313) x BE599526 (1-487)

QY 52 GlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCys 71

Db 203 CAGAGGATGGGGCCCTCTGCTCAGGCTCTTCACAGACTGCTTGTTCAGAGTTGC 262

QY 72 AspGlySerIleLeuLeuAspAlaGlyGlyLysThrAlaGlyProasn 88

Db 263 GACGGCTCGATCTTCTTCAGCCCGAGGGGAGAGACGGCCGCCAC 313

RESULT 14

BE600529 495 bp mRNA linear EST 18-AUG-2000

LOCUS P11.94.C12.b1.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,

DEFINITION mRNA sequence.

ACCESSION BE600529

VERSION BE600529.1 GI:9855602

KEYWORDS EST.

SOURCE

ORGANISM

Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoidae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 495)

AUTHORS Cordomier-Pratt, M.-M., Gingle, A., Dean, R., Sudman, M. and Pratt

TITLE An EST database from Sorghum: pathogen-induced plants

JOURNAL Unpublished (2000)

JOURNAL  
COMMENT

Unpublished (2000)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude polyA, vector and regions  
below phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: JEN REV  
High quality sequence stop: 481  
POLYA-No.

FEATURES  
source

Location/Qualifiers  
1..495

/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Pathogen induced 1 (P11)"  
/note="Organ: Anthracnose-infected leaves from  
two-week-old sorghum plants 48 hr after inoculation;  
Vector: pBluescript II from lambda Zap II; Site.1: XhoI;  
Site.2: EcoRI; Two-week-old sorghum plants (B7X 623  
cultivar) were infected with pathogen (isolate FRM421 of  
Colletotrichum graminicola, which is a sorghum isolate).  
RNA was prepared from infected leaves harvested from 45  
seedlings 48 hours after inoculation. Note: young  
seedlings (2 weeks old) exhibit juvenile resistant  
reaction, which is an incompatible interaction. As they  
grow older (4 weeks or older), plants resume susceptibility  
to anthracnose disease. The library was made from poly-A  
RNA in the cloning vector lambda Zap II. Clones to be  
sequenced were prepared by mass excision. WARNING: While  
most or all ESTs are expected to derive from the host  
plant, no effort was made to eliminate ESTs deriving from  
the pathogen."

BASE COUNT 82 a 170 c 145 g 97 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 2,02e-25 Length: 495  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.82% Indels: 0  
DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE600529 (1-495)

QY 52 GlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCys 71  
|||||  
Db 208 CAGAGGATGGGCGCTCTGCTCAGGCTCTTCCACAGACTGCTTTTCAAGGTTGC 267

QY 72 AspGlySerIleLeuLeuAspAlaGlyGlnGlyThrAlaGlyProAsn 88  
|||||  
Db 268 GACGGCTCGATCTTCTCGACGCGGAGGAGGAGACGCGGCCCAAC 318

RESULT 15  
BE594365 522 bp mRNA linear EST 18-AUG-2000  
LOCUS P11\_32\_E10\_b2\_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,  
DEFINITION mRNA sequence.  
ACCESSION BE594365  
VERSION BE594365.1 GI:9849438  
KEYWORDS EST.  
SOURCE sorghum,  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 522)  
Cordonnier-Pratt, M.-M., Gingle, A., Dean, R., Sudman, M. and Pratt,  
L.H.

TITLE  
JOURNAL  
COMMENT

An EST database from Sorghum: pathogen-induced plants  
Unpublished (2000)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude polyA, vector and regions  
below phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: JEN REV  
High quality sequence stop: 475  
POLYA-No.

FEATURES  
source

Location/Qualifiers  
1..522

/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Pathogen induced 1 (P11)"  
/note="Organ: Anthracnose-infected leaves from  
two-week-old sorghum plants 48 hr after inoculation;  
Vector: pBluescript II from lambda Zap II; Site.1: XhoI;  
Site.2: EcoRI; Two-week-old sorghum plants (B7X 623  
cultivar) were infected with pathogen (isolate FRM421 of  
Colletotrichum graminicola, which is a sorghum isolate).  
RNA was prepared from infected leaves harvested from 45  
seedlings 48 hours after inoculation. Note: young  
seedlings (2 weeks old) exhibit juvenile resistant  
reaction, which is an incompatible interaction. As they  
grow older (4 weeks or older), plants resume susceptibility  
to anthracnose disease. The library was made from poly-A  
RNA in the cloning vector lambda Zap II. Clones to be  
sequenced were prepared by mass excision. WARNING: While  
most or all ESTs are expected to derive from the host  
plant, no effort was made to eliminate ESTs deriving from  
the pathogen."

BASE COUNT 88 a 182 c 151 g 101 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,14e-25 Length: 522  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.82% Indels: 0  
DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE594365 (1-522)

QY 52 GlnArgMetGlyAlaSerLeuLeuArgLeuPhePheHisAspCysPheValGlnGlyCys 71  
|||||  
Db 208 CAGAGGATGGGCGCTCTGCTCAGGCTCTTCCACAGACTGCTTTTCAAGGTTGC 267

QY 72 AspGlySerIleLeuLeuAspAlaGlyGlnGlyThrAlaGlyProAsn 88  
|||||  
Db 268 GACGGCTCGATCTTCTCGACGCGGAGGAGGAGACGCGGCCCAAC 318

Search completed: May 3, 2003, 14:52:12  
Job time: 1086 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 3, 2003, 12:06:43 ; Search time 1082 Seconds

(without alignments)  
4685.016 Million cell updates/sec

Title: US-10-047-825-4

Perfect score: 1604

Sequence: 1 MASPTMQCLVAVSLSCVA.....YGVLTGTAQGRRCRCRVVNS 313

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame-p2n.model -DEV=rlh  
-O=/cgn2\_1/USFTO.spool/US10047825/runat.26042003.111958.7143/app.query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human4.0.cdi -LIST=45  
-DOCALLIGN=200 -THR=SCORE-pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pio -NOR=ext -HEADS=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US10047825 @cgn2\_1.1.1456 @runat.26042003.111958.7143 -NCPU=6 -ICPU=3  
-NO\_XLPEXT -NO\_MMAP -LARGEJURY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	999.5	62.3	848	9	AA979912
2	964	60.1	633	10	AA671673
3	942	58.7	1303	11	AY107804
4	928.5	57.9	602	12	BG463106
5	914	57.0	625	10	AA672124
6	904.5	56.4	593	12	BG462797
7	891	55.5	593	10	BE362200
8	879	54.8	591	10	BE362356
9	871	54.3	615	10	BE361528
10	841.5	52.5	1570	11	AY108658
11	838	52.2	829	12	BE263468
12	822	51.2	1275	11	AY105947
13	809	50.4	589	10	BE597430
14	793	49.4	561	10	BE361738
15	793	49.4	570	10	BE358471
16	788	49.1	555	10	AA671012
17	788	49.1	564	10	BE600142
18	782	48.8	560	10	BE600110
19	781	48.7	566	10	BE366224
20	774	48.3	533	10	BE125306
21	771	48.1	569	10	BE357554
22	769	47.9	566	10	BE600823
23	767.5	47.8	1346	11	AY105520
24	764	47.6	815	12	BE267871
25	755.5	47.1	874	13	BI952385
26	750	46.8	744	14	BQ744368
27	749	46.7	1504	11	AY109094
28	743.5	46.4	983	10	BE034991
29	743	46.3	537	10	BE595535
30	742	46.3	446	12	BG841858
31	740	46.1	524	10	BE597461
32	739.5	46.1	798	14	BQ164905
33	736.5	45.9	813	13	BI952379
34	736	45.9	522	10	BE594365
35	725	45.2	834	12	BE267576
36	724.5	45.2	845	13	BM359358
37	722	45.0	641	10	BE555297
38	718.5	44.8	925	12	BG442996
39	716	44.6	672	10	AV925367
40	708	44.1	779	14	BQ165051
41	708	44.1	1231	11	AY108615
42	705	44.0	803	12	BE267910
43	703.5	43.9	930	10	BE412497
44	702	43.8	762	12	BG581250
45	700	43.6	495	10	BE600529

#### ALIGNMENTS

RESULT 1  
AA979912  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AA979912 848 bp mRNA linear EST 26-MAY-1998  
MEST3-D6.TW1412.seg ISUM2 Zea mays cDNA clone MEST3-D6 5', mRNA  
AA979912  
AA979912.1 GI:3157290  
EST.  
Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 848)  
Wen,T.J., Ashlock,D.A. and Schnable,P.S.

TITLE Expressed Sequence Tags from B73 Maize Seedlings  
JOURNAL Unpublished (1997)  
COMMENT Contact: Schnable, PS  
Schnable Laboratory  
Iowa State University  
G405 Agronomy, Ames, IA 50011, USA  
Tel: (515)-294-0975  
Fax: (515)-294-2299  
Email: schnable@iastate.edu

PCR Primers  
FORWARD: tw1412 (5'-GAAGATACCCGACCAACC-3')  
BACKWARD: tw1412 (5'-TAATACACTCATATAGGC-3')  
Plate: MEST3 row: D column: 6  
Seq primer: tw1412 (5'-GAAGATACCCGACCAACC-3').

FEATURES  
source  
1. 848  
/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ISUM2"  
/tissue\_type="above ground tissues"  
/dev\_stage="two-leaf-stage green seedling"  
/lab\_host="XLI-MFR Blue"  
/note="Organ: green seedlings; Vector: PAD-GAL4; Site:1:  
EcoRI; Site:2: XhoI; ds-cDNA molecules were generated as  
follows. First-strand cDNA was prepared from oligo-dT  
selected mRNA by priming with an XhoI oligo-dT primer. The  
resulting DNA:RNA hybrid was treated with RNase H and used  
as a template for DNA Pol-catalyzed second strand  
synthesis. After the addition of EcoRI adaptors, the  
ds-cDNAs were digested with XhoI and size-selected. The  
resulting molecules were directionally cloned into the  
EcoRI and XhoI sites of the HybridZAP lambda vector  
(Stratagene) and excised as PAD-GAL4 phagemids."

BASE COUNT 162 a 284 c 236 g 147 t 19 others  
ORIGIN

Alignment Scores:  
Pred. No.: 8.49e-93 Length: 848  
Score: 999.50 Matches: 204  
Percent Similarity: 92.38% Conservatave: 2  
Best Local Similarity: 91.48% Mismatches: 17  
Query Match: 62.31% Indels: 2  
DB: 9 Gaps: 0

US-10-047-825-4 (1-313) x AA979912 (1-848)

QY 87 ProAsnLeuAsnSerValAlaGlyPheGluValIleAspThrIleTyrArgAsnValGlu 106  
Db 1 CCGAACCTGAACCTGGCGCGCTTGAAGTCATCGACACCAAGCGAAGCGTCGAG 60

QY 107 AlaAlaCysProGlyValValSerCysAlaAspIleLeuAlaLeuAlaAlaArgAspGly 126  
Db 61 GCGCGGCGCGCGCGCGCTGCTGCGCGGACATCTCGCGCTTCCCGCGCGCGGA 120

QY 127 ThrAsnLeuLeuGlyGlyProThrTrpSerValProLeuGlyValArgAspSerThrThr 146  
Db 121 ACCAACCTTCTCGCGCGCGCGACCTGAGCGTGCCTCGCGCGCGCGGACTCGAGAG 180

QY 147 AlaSerAlaSerLeuAlaAsnSerAsnProProProThrAlaSerLeuGlyThrLeu 166  
Db 181 GCCAGCGCGCTGCTGCGCAACACACCTCCGCCCGGACGCGCGCTCGGACGCTC 240

QY 167 IleSerLeuPheGlyArgGlnGlyLeuSerProArgAspMetThrAlaLeuSerGlyAla 186  
Db 241 ATCTCCCTGTTGGGACGAGCGGCGCTGCGCGGACATAGCGCGCTGTCGGCGCG 300

QY 187 HisThrIleGlyAlaAlaArgCysThrThrPheArgGlyValGlyIleTyrGlyAspThrAsp 206  
Db 301 CACACCATCGGCGAGCGCGCTGACACCATTCGCGCGCGCATCTACGCGACACGAGC 360

QY 207 IleAsnAlaSerPheAlaAlaLeuArgGlnGlnThrCysProArgSerGlyLeuAspGly 226

Db 361 ATCAACGCCCTCTCTCGCGCGCTGCGCGACGAGCTCCCGGCTCGCGCGGACGCGC 420  
QY 227 AsnLeuAlaProIleAspValGlnThrProValArgPheAspThrAlaTyrPheThrAsn 246  
Db 421 AACCTGGCGCCATCGAGCTGAGCGCGGAGGTTGACACGCGCTTCTTACACAC 480

QY 247 LeuLeuSerArgArgGlyLeuPheHisSerAspGlnGluLeuPheAsnGlyGlySerGln 266  
Db 481 CAGCTGTGGCGGANGGCGCTGTTCACATCGGACGACGACGACGCTTCAACNGCGGTCGCA 540

QY 267 AspAlaLeuValArgGlnTyrSerAlaSerAlaSerLeuPheAsnAlaAspPheValAla 286  
Db 541 GACGCGCTGCTGCAACCAATACAGCGCGCGCTTCTTCAACGCGGACTT-GTGCA 599

QY 287 AlaMetIleArgMetGlyAsnValGlyValLeuThrGlyThrAlaGlyGlnIleArgArg 306  
Db 600 ACCATGATTANGATGGGAACCTTTGGTGTCTCAACCGGACC-GTTGGNAGATCAAGGCA 658

QY 307 AsnCysArg 309  
Db 659 ACTGCGCG 667

RESULT 2  
AM671673 633 bp mRNA linear EST 19-JUL-2000  
LOCUS  
DEFINITION LG1\_349\_F09\_b1\_A002 Light Grown 1 (LG1) sorghum bicolor cDNA, mRNA  
sequence.  
ACCESSION AM671673  
VERSION AM671673.1 GI:7535567  
KEYWORDS EST.  
SOURCE  
ORGANISM  
Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.  
1 (bases 1 to 633)  
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.  
An EST database from Sorghum: light-grown seedlings  
Unpublished (2000)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mpratt@uga.edu  
Sequences have been trimmed to exclude POLYA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: JEN REV  
High quality sequence stop: 576  
POLYA-No.

FEATURES  
source  
1. 633  
/location/Qualifiers  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Light Grown 1 (LG1)"  
/note="Organ: 10- to 14-day-old light-grown (greenhouse)  
seedlings; Vector: lambda Zap; Site:1: XhoI; Site:2: EcoRI  
; The library was made from poly-A RNA in the cloning  
vector lambda Zap II. Clones to be sequenced were  
prepared by mass excision."

BASE COUNT 105 a 230 c 190 g 108 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.61e-89 Length: 633  
Score: 964.00 Matches: 188  
Percent Similarity: 94.23% Conservatave: 8  
Best Local Similarity: 90.38% Mismatches: 12  
Query Match: 60.10% Indels: 0  
DB: 10 Gaps: 0



```

      833 CACCACACCGCTTACAGCTTGCACACGCTTACACCAACCTCTAGCCAGAGG 892
      252 YleuphehisSerAspGlnGluLeuPhehisnglySerGlnAspAlaLeuValArg1 272
      893 GCTCTGACACTCGACAGAGAGCTCTTCAACGCGGACAGCAGACCAACACCTCAGGAA 952
      272 nTYrSerAlaSerAlaSerLeuPhehisnAlaAspPheValAlaAlaMetIleArgMet1 292
      953 CTTCGGCTCCACACTCGCGCCCTTCAGACAGCCCTTCGCGCGGCGCATGTGTAAGATGSG 1012
      292 YasnValGlyValLeuThrglyThzAlaGlyGlnIleArgAsnGlyArgValValns 312
      1013 CAACCTCAGCCCGCTCAGGATCTCAGGCGCAGATCAGGCTTACCTGCTCCACAGTGAA 1072
      312 n 312
      1073 C 1073

RESULT 4
LOCUS BG463106 602 bp mRNA linear EST 20-MAR-2001
DEFINITION EMI_47_E12_b1_A002 Embryo 1 (EMI) Sorghum bicolor cDNA, mRNA
ACCESSION BG463106
VERSION BG463106.1 GI:13390208
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 602)
REFERENCE Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.
AUTHORS An EST database from Sorghum: developing embryos
TITLE Unpublished (2000)
JOURNAL Contact: Cordonnier-Pratt MM
COMMENT Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpatrteuga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 548
POLYA-NO.

FEATURES
source location/Qualifiers
1..602
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Embryo 1 (EMI)"
/note="Organ: Embryos germinated for 24 hr; Vector:
pBluescript II from lambda zap II; Site_1: xhoI; Site_2:
BcoRI; The library was made from polyA RNA in the cloning
vector lambda zap II. Clones to be sequenced were
prepared by mass excision."
BASE COUNT 101 a 219 c 192 g 90 t
ORIGIN

Alignment Scores:
Pred. No.: 1,12e-85 Length: 602
Score: 928.50 Matches: 180
Percent Similarity: 93.94% Conservative: 6
Best Local Similarity: 90.91% Mismatches: 11
Query Match: 57.89% Indels: 1
DB: 12 Gaps: 1

US-10-047-825-4 (1-313) x BG463106 (1-602)
QY 59 LeuArgLeuPhehisSerPheValGlnGlyCysAspGlySerIleLeuLeuasp 78

```

```

      9 CTACAGGCTCTTCTCCACAGACTGCTTGTCAAGGTTCGACGCGCTGATCTTCGAC 68
      79 AlaGlyGlyGlyTyrThrAlaGlyProAsnLeuInsSerValArgGlyPheGluValIle 98
      69 GCGGAGGGGAGAGACGCGCGCGCCCAACGCCCAATTCGGCGCGCTTGAAGTATC 128
      99 AspThrIleTyrArgAsnValGluAlaAlaCysProGlyValValSerCysAlaAspIle 118
      129 GACACCATCAGAACCAACAGGTGAGAGCCGCGTCCCGCGCGTGTGCGCCGACATC 188
      119 LeuAlaLeuAlaAlaArgAspGlyThrAsnLeuLeuGlyGlyProThrTyrPseValPro 138
      189 CTTCGCCCTTGGCGCGCGAGACCAACATGCTCGCGCGCGGACCTGACACGTCGCG 248
      139 LeuGlyArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAspProPro 158
      249 CTGGGCGCGCGGACTCGACAGACGCGCGCTTCCTTCCCAACACACACCTCCCGCAG 308
      159 ProThrAlaSerLeuGlyThrLeuIleSerLeuPheGlyArgGlnGlyLeuSerProArg 178
      309 TCGACGGCGAGCGCTGGGTAACGCTCATCTCGCTTCAGACAGGCGAGGCGCTGGCGCGC 368
      179 AspMetThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArgCysThrThrPheArg 198
      369 GACATGACGGCGCTGTGCGCGCGCACACCATCGGCGAGCGCGCTCACACCTTCGCG 428
      199 GlyArgIleTyrGlyAspThrAspIleAsnAlaSerPhe--AlaAlaLeuArgGlnGln 217
      429 AGCCGCATCTACGGCGACACCAACATCAGCCCTCTTCGCGCGCGCGCTCCGACAG 488
      218 ThrCysProArgSerGlyGlyAspGlyAsnLeuAlaProIleAspValGln;nrProVal 237
      489 ACGTTCGCCGAGCTCGCGCGCGAGCGCAACTGTGCGCCCATGAGCGAGACGCCGACG 548
      238 ArgPheAspThrAlaTyrPheThrAsnLeuSerArgArgGlyLeuPheHis 255
      549 AGTTTCGACAGCACTACTACACCAACCTGTGTGCAAGGCGCGCTTCCAC 602

RESULT 5
LOCUS AM672124 625 bp mRNA linear EST 19-JUL-2000
DEFINITION LG1_357_F02_b1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
ACCESSION AM672124
VERSION AM672124.1 GI:7536032
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 625)
REFERENCE Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
AUTHORS An EST database from Sorghum: light-grown seedlings
TITLE Unpublished (2000)
JOURNAL Contact: Cordonnier-Pratt MM
COMMENT Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpatrteuga.edu
Sequences have been trimmed to exclude polyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 508
POLYA-NO.

FEATURES
source location/Qualifiers
1..625
/organism="Sorghum bicolor"
/db_xref="taxon:4558"

```



Db 489 ACGTCCCGGAGTCCGGGGGAGGAGCACTTGGCGCCCATGGAGCTGACAGCCGACG 548

QY 238 ArpheasphrathatyrpethrAsnleuSerArgArg 251  
 |||||  
 |||||

Db 549 AGTTTCGACAGGACTACTACCAACTCTGTTCGACAGG 590

RESULT 7  
 BE362200 593 bp mRNA linear EST 20-JUL-2000  
 LOCUS DGI\_85\_C06\_b1\_A002 Dark grown 1 (DGI) Sorghum bicolor cDNA, mRNA  
 DEFINITION sequence.  
 ACCESSION BE362200  
 VERSION BE362200.1 GI:9303757  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 593)  
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 ,L.H.  
 TITLE An EST database from Sorghum: dark-grown seedlings  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@pratt.uga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: JEN REV  
 High quality sequence stop: 523  
 POLYA-NO.

FEATURES  
 source location/Qualifiers  
 1..593 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Dark Grown 1 (DGI)"  
 /note="Organ: 5-day-old dark-grown seedlings; Vector:  
 Lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was  
 made from poly-A RNA in the cloning vector Lambda Zap II.  
 Clones to be sequenced were prepared by mass excision."  
 BASE COUNT 94 a 210 c 184 g 105 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8.09e-82 Length: 593  
 Score: 891.00 Matches: 175  
 Percent Similarity: 94.308 Conservative: 7  
 Best Local Similarity: 90.67 Mismatches: 11  
 Query Match: 55.55% Indels: 0  
 DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE362200 (1-593)

QY 1 MetAlaSerProThrLeuMetGlnCysLeuValAlaValSerLeuLeuSerCysValAla 20  
 |||||  
 |||||

Db 15 ATGGCGGCTCCATCTTGATGCAATGCTTGCGCATTTCCCTCTCTCTGTCGCGC 74  
 |||||

QY 21 HisAlaGlnLeuSerProThrPheTyrAlaSerSerCysProAsnLeuGlnSerIleVal 40  
 |||||  
 |||||

Db 75 CAGCGACACTCTCGACCGCGTTCTATGATCATCTCTCGCCCAACCTGCAGACATCGTG 134  
 |||||

QY 41 ArgAlaAlaMetThrGlnAlaValAlaSerGlnInaArgMetGlyAlaSerLeuLeuArg 60  
 |||||  
 |||||

Db 135 CCGAGGCGCATGATGACGCGCTCTCAGCAATGACCGAGAGTGCGCGCTCTCTGTCAGG 194  
 |||||

QY 61 LeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGly 80  
 |||||

Db 195 CTCTCTTCACAGACTCGTTTGTCAGGTTGCGAGCGCTCATCTTCTCGACGCCGA 254

QY 81 GlyGluLysThrAlaGlyProAsnLeuAsnSerValArgGlyPheGluValIleAspThr 100  
 |||||  
 |||||

Db 255 GGGGAGACACGCGCGCGCCCAACGCGCAATTTGGCGCGCGCTTGAGGTATCGACACG 314

QY 101 IleIysArgAsnValGlnAlaIleAcysProGlyValValSerCysAlaAspIleLeuAla 120  
 |||||  
 |||||

Db 315 ATCAGACACCAAGCTGAGAGCGCGCGCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 374

QY 121 LeuAlaAlaArgAspGlyThrAsnLeuGlnGlyGlyProThrTrpSerValProLeuGly 140  
 |||||  
 |||||

Db 375 CTTCGCGCGCGAGACGAGCAACATCTGCTGCGCGCGCGACCTGAGAGTGGCGCGGCG 434

QY 141 ArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAspProProProThr 160  
 |||||  
 |||||

Db 435 CCGCGGACTCGACGACGCGCGCGCGCTCTGTCACAGACACCTCCGACATCGACG 494

QY 161 AlaSerLeuGlyThrLeuLeuSerLeuPheGlyArgGlnGlyLeuSerProArgAspMet 180  
 |||||  
 |||||

Db 495 GCGAGCTGGTACGCTATCTCGCTGTTCAAGACAGGAGGCGCTGTGCGCGCACATG 554

QY 181 ThrAlaLeuSerGlyAlaHisThrIleGlyGlnAlaArg 193  
 |||||  
 |||||

Db 555 ACGGCGCTGTGCGCGCGCGCACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 593

RESULT 8  
 BE362356 591 bp mRNA linear EST 20-JUL-2000  
 LOCUS DGI\_86\_C05\_b1\_A002 Dark grown 1 (DGI) Sorghum bicolor cDNA, mRNA  
 DEFINITION sequence.  
 ACCESSION BE362356  
 VERSION BE362356.1 GI:9303913  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 591)  
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 ,L.H.  
 TITLE An EST database from Sorghum: dark-grown seedlings  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@pratt.uga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: JEN REV  
 High quality sequence stop: 531  
 POLYA-NO.

FEATURES  
 source location/Qualifiers  
 1..591 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Dark Grown 1 (DGI)"  
 /note="Organ: 5-day-old dark-grown seedlings; Vector:  
 Lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was  
 made from poly-A RNA in the cloning vector Lambda Zap II.  
 Clones to be sequenced were prepared by mass excision."  
 BASE COUNT 94 a 210 c 181 g 105 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.39e-80 Length: 591  
 Score: 879.00 Matches: 173  
 Percent Similarity: 93.75% Conservative: 7

Best Local Similarity: 90.10% Mismatches: 12  
 Query Match: 54.80% Indels: 0  
 DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE361528 (1-591)

```

Oy 1 MetalaserProthrlleuMetGlnCysLeuValAlaValSerleuLeuSerCysValAla 20
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 15 ATGGCGGCTCTACCTTATGCAATGCTTGGTCCGCAATTCCTCTCTCTGTCGCC 74
Oy 21 HislagnleuSerProthrlPheTyrlAlaserSerCysProasleuGlnSerlleval 40
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 75 CACGCACAGCTCTCGACACAGCTTATGCAATCTCTCCACCACTCGACGATCGTG 134
Oy 41 ArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGlyAlaSerleuLeuArg 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 135 CGGAGGCGGATGATCCAGGCCCTCGCAATGACGAGAGAGGCGGCTCTCTCTCAGG 194
Oy 61 LeuphepHeHIsaSPCySPheValGlnGlyCysAspGlySerlleLeuLeuAspAlaGly 80
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 195 CTCTTCTTCCACGACTGCTTCTTCAAGTTGCGAGCGCTCGATCTCTCGACGCCGA 254
Oy 81 GlyGlnuYsthrAlaGlyProasleuAsnSerValArgGlyPheGluValIleAspThr 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 255 GGGGAGAGAGACGGCGGCCCAACCCCAATTCGGCGCGCTTGAAGTCATCGACACC 314
Oy 101 IleLysArgAsnValGlnAlaAlaCysProGlyValAlaSerCysAlaAspIleLeuAla 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 315 ATCAAGACCAACAGTGAGGCGCGCGGCCGCGTGGTGCGTGGCGACATCTTCTGCC 374
Oy 121 LeuAlaAlaArgAspGlyThrAsnleuLeuGlyGlyProthrlTrpSerValProleuGly 140
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 375 CTGGCGCGCGAGAGGAACCAATCTGCTCGCGCGCGAGCTGGAAGTGGCGGTGGGC 434
Oy 141 ArgAlaGAspSerThrThrAlaserAlaserleuAlaAsnSerAsnProProProthrlr 160
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 435 CGCGCGGACTCGACACAGCGCGCGCTCCCTTGGCAACAGCAACCTCCGCAAGTCGACG 494
Oy 161 AlaserleuGlyThrleuIleSerleuPheGlyArgGlnGlyLeuSerProArgAspMet 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 495 GGGAGGCTGNGTACGCTCATCTCGCTGTTACAGCAGGCGAGGCGCTTGGCGCGGACATG 554
Oy 181 ThrAlaLeuSerGlyAlaHisThrIleGlyGlnAla 192
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 555 ACGGCGCTGTCGGGCGCGCACACATCGGCGAGGCC 590

```

RESULT 9 BE361528 615 bp mRNA linear EST 20-JUL-2000  
 BE361528  
 LOCUS DGI\_81.G11.A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA  
 DEFINITION  
 BE361528  
 ACCESSION BE361528  
 VERSION BE361528.1 GI:9303155  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Spumetocyt: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Sorghum.  
 REFERENCE 1 (bases 1 to 615)  
 AUTHORS Cordomier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt  
 L.H.  
 TITLE An EST database from Sorghum: dark-grown seedlings  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordomier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@pratt.uga.edu

Sequences have been trimmed to exclude polyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence

1s 20.  
 Seq primer: JEN REV  
 High quality sequence stop: 527  
 POLY-A-NO.

FEATURES  
 source Location/Qualifiers  
 1. 615

/organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone.lib="Dark Grown 1 (DGI)"  
 /note="Organ: 5-day-old dark grown seedlings; Vector:  
 Lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was  
 made from poly-A RNA in the cloning vector Lambda Zap II.  
 Clones to be sequenced were prepared by mass excision."

BASE COUNT 101 a 215 c 186 g 113 t

Alignment Scores:  
 Pred. No.: 9.84e-80 Length: 615  
 Score: 871.00 Matches: 171  
 Percent Similarity: 94.18% Conservative: 7  
 Best Local Similarity: 90.48% Mismatches: 11  
 Query Match: 54.30% Indels: 0  
 DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE361528 (1-615)

```

Oy 1 MetalaserProthrlleuMetGlnCysLeuValAlaValSerleuLeuSerCysValAla 20
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 48 ATGGCGGCTCTACCTTATGCAATGCTTGGTCCGCAATTCCTCTCTCTGTCGCC 107
Oy 21 HislagnleuSerProthrlPheTyrlAlaserSerCysProasleuGlnSerlleval 40
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 108 CACGCACAGCTCTCGACACAGCTTATGCAATCTCTCCACCACTCGACGATCGTG 167
Oy 41 ArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGlyAlaSerleuLeuArg 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 168 CGGAGGCGGATGATCCAGGCCCTCGCAATGACGAGAGATGGCGGCTCTCTGCTCAGG 227
Oy 61 LeuphepHeHIsaSPCySPheValGlnGlyCysAspGlySerlleLeuLeuAspAlaGly 80
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 168 CGGAGGCGGATGATCCAGGCCCTCGCAATGACGAGAGATGGCGGCTCTCTGCTCAGG 227
Oy 81 GlyGlnuYsthrAlaGlyProasleuAsnSerValArgGlyPheGluValIleAspThr 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 228 CTCTTCTTCCACGACTGCTTCTTCAAGTTGCGAGCGCTCGATCTCTCGACGCCGA 287
Oy 101 IleLysArgAsnValGlnAlaAlaCysProGlyValAlaSerCysAlaAspIleLeuAla 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 348 ATCAAGACCAACAGTGAGGCGCGCGGCCGCGTGGTGCGTGGCGACATCTTCTGCC 407
Oy 121 LeuAlaAlaArgAspGlyThrAsnleuLeuGlyGlyProthrlTrpSerValProleuGly 140
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 408 CTGGCGCGCGAGAGGAACCAATCTGCTCGCGCGCGAGCTGGAAGTGGCGGTGGGC 467
Oy 141 ArgAlaGAspSerThrThrAlaserAlaserleuAlaAsnSerAsnProProProthrlr 160
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 468 CGGCGGACTCGACACAGCGCGCGCTCCCTTGGCAACAGCAACCTCCGCAAGTCGACG 527
Oy 161 AlaserleuGlyThrleuIleSerleuPheGlyArgGlnGlyLeuSerProArgAspMet 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 528 GGGAGGCTGNGTACGCTCATCTCGCTGTTACAGCAGGCGAGGCGCTTGGCGCGGACATG 587
Oy 181 ThrAlaLeuSerGlyAlaHisThrIle 189
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 588 ACGGCGCTGTCGGGCGCGCACACCATC 614

```

RESULT 10 AY108658 1570 bp mRNA linear HTC 25-MAY-2002  
 AY108658  
 LOCUS AY108658  
 DEFINITION zea mays PC0073925 mRNA sequence.  
 ACCESSION AY108658  
 VERSION AY108658.1 GI:21211830  
 KEYWORDS HTC.



SOURCE	ORGANISM
Zea mays.	Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.	
Hanley,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whiltsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M., and Tinsley,S.V.	
Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes	
Unpublished (2002)	
2 (bases 1 to 1570)	
Coe,E.C.	
Direct Submission	
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA	
Location/Qualifiers	
source	
/organism="Zea mays"	
/db.xref="MaizEDB:634070"	
/db.xref="taxon:4577"	
/clone="PCO073925"	
/clone_id="Maize Mapping Project/DuPont Consensus library"	
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACS in conjunction with the Maize Mapping Project"	
BASE COUNT	373 a 460 c 414 g 323 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3,97e-76
Score:	841.50
Percent Similarity:	61.35%
Best Local Similarity:	47.30%
Mismatches:	52
Indels:	63
Gaps:	5
US-10-047-825-4 (1-313) x AY108658 (1-1570)	
OY	5 ThrlleuwtgncysleuValAlAvalSerLeuSerCySValAlAhislaglnleu 24
Db	126 AGCTGTGGTGGTGCGGCGCGCGCTGTGCAGCAACGCGCCGACGACTG 185
OY	25 SerProThrPheTyrrAlaSerSerCySProAsnLeuGlSerlleValArgAlaAmet 44
Db	186 TCgtTCGAAGCTTCAGACAGCTGTGCCCGCATGTGCACCAATCACAGCGGCGTg 245
OY	45 ThrGlnAlaValAlaSerGlnArgMetGlyAlaSerLeuLeuArgLeuphePhenIs 64
Db	246 AACTCCCGCGGGCGAGAGCTCGGGTGGGGCGGTGTGCGGCTCCACTTCACC 305
OY	65 AspCysPheVal
Db	306 GACTGCTTCGTCACAAGCAAGCTAGCTGTCTCAGATGATCTATCTACTATATAT 365
OY	68
Db	366 AAGCATGATTTCCTTTCTACTAGCTTAGCATCTCGTCATTTTTAATTGAAGTAAAG 425
OY	68
Db	426 ATTAGACGTCGTATATGATGCTAATTAAACAGAGGATCATGTGTAATTTCTG 485
OY	69 -----GlnIcylCysAspGlySerlleleuLeu---AspAlaGlylGylulSthr 84
Db	486 GTGGTCACAGAGGGCGCGACGCTGCTGTCGACAGCGTCCGGGAGACAGACC 545
OY	85 AlagIyProasnLeuasn---SerValArgGlyPheGluValIlleasphrlllelsarg 103





D5 851 AACCTCTGGCCGCAAGGSCCTTCTAAGCTCTGATGAGATTCTGTTAACCAAGAGCGCT 910

Best Local Similarity: 89.39% Mismatches: 12

Percent Similarity:	93.30%	Conservative:	7
Best Local Similarity:	89.39%	Mismatches:	12

Query Match: 50.448 Indels: 0  
DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE597430 (1-589)

QY 1 MetLaserProthrlMetGlnCysLeuValAlaValSerLeuLeuSerCysValAla 20  
DB 52 ATGGCGGCTCTTACCTTGATGCAATGCTTGCTCCCATTTCCCTCTCTGCTGCTGCC 111

QY 21 HisLagInLeuSerProthrlPheTyrAlaSerSerCysProsnLeuGlnSerLleVal 40  
DB 112 CACCCACAGCTCTGACACAGCTTCTATGCAATCTCTGCCCCAACCTGACAGATCTGTG 171

QY 41 ArgAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGlyAlaSerLeuLeuArg 60  
DB 172 CGGAGGGGATGATCCAGCGCCCTCAGCAATGACACAGAGATGGCGGCTCTGCTGCTAGG 231

QY 61 LeuPhePheHisAspCysPheValGlnGlyCysAspGlySerLleLeuLeuAspAlaGly 80  
DB 232 CTCTCTTCCACGACTGCTTGTTCAGGTTGCGACGCGCTCGAATCTTCTGACGCGCGGA 291

QY 81 GlyLulysThrAlaGlyProsnLeuAsnSerValArgGlyPheGluValLleAspThr 100  
DB 292 GGGAGAAAGACGGCCGCCCCCAACCAATTCGCGCGCTTGAGGTATCGACACC 351

QY 101 IleLysArgAsnValGluAlaAlaCysProGlyValValSerCysAlaAspLleLeuAla 120  
DB 352 ATCAAGACCAACGTGAGGCGCGCTGCCCGGCTGTGCGTGGCGGACATCTTGCC 411

QY 121 LeuAlaAlaArgAspGlyThrAsnLeuLeuGlnGlyLysProThrTrpSerValProLeuGly 140  
DB 412 CTTCCTCCGCGAGACGAGACCAATCTGCTGCGCGCGCCAGCTGNAACGTGCGCTGGCG 471

QY 141 ArgArgAspSerThrThrAlaSerAlaSerLleuAlaAsnSerAsnProProProThr 160  
DB 472 CGGGGGAGCTGACAGAGCGGAGCGCTCCCTTCCCAACGACCACTCCCGCAGTCGACG 531

QY 161 AlaSerLeuGlyThrLeuLleSerLeuPheGlyArgGlnGlyLeuSerProArgAsp 179  
DB 532 GCGAGCCTGGGTACGCTCATCTGCTGTGACGAGGAGCGCTGTGGCGCGGAC 588

RESULT 14  
LOCUS BE361738 561 bp mRNA linear EST 20-JUL-2000  
DEFINITION DGI\_82\_H08\_b1\_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA  
ACCESSION BE361738  
VERSION BE361738.1 GI:9303295  
KEYWORDS EST.  
SOURCE sorghum.  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.  
REFERENCE 1 (bases 1 to 561)  
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt,  
L.H.  
TITLE An EST database from Sorghum: dark-grown seedlings  
JOURNAL Unpublished (2000)  
COMMENT Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@pratt.uga.edu  
Sequences have been trimmed to exclude polyA, vector and regions  
below phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: JEN REV  
High quality sequence stop: 514  
POLYA-No.  
FEATURES Location/Qualifiers

source 1..561  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Dark Grown 1 (DGI)"  
/note="Organ: 5-day-old dark-grown seedlings; Vector:  
lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was  
made from poly-A RNA in the cloning vector lambda Zap II.  
Clones to be sequenced were prepared by mass excision."

BASE COUNT 94 a 195 c 166 g 106 t

ORIGIN

Alignment Scores:  
Pred. No.: 9.66e-72 Length: 561  
Score: 793.00 Matches: 155  
Percent Similarity: 94.19% Conservative: 7  
Best Local Similarity: 90.12% Mismatches: 10  
Query Match: 49.44% Indels: 0  
DB: 10 Gaps: 0

US-10-047-825-4 (1-313) x BE361738 (1-561)

QY 1 MetLaserProthrlMetGlnCysLeuValAlaValSerLeuLeuSerCysValAla 20  
DB 45 ATGGCGGCTCTTACCTTGATGCAATGCTTGCTCCCATTTCCCTCTCTGCTGCTGCC 104

QY 21 HisLagInLeuSerProthrlPheTyrAlaSerSerCysProsnLeuGlnSerLleVal 40  
DB 105 CACGCACAGCTCTGACACGCGCTTATGCAATCTCTGCCCCAACCTGACAGATCTGTG 164

QY 41 ArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGlyAlaSerLeuLeuArg 60  
DB 165 CGGAGGCGATGATCCAGGCGCTTATGCAATGACAGAGATGGCGGCTCTCTGCTGAGG 224

QY 61 LeuPhePheHisAspCysPheValGlnGlyCysAspGlySerLleLeuLeuAspAlaGly 80  
DB 225 CTCTCTTCCACGACTGCTTGTTCAGGTTGCGACGCGCTGATCTTCTGACGCGCGGA 284

QY 81 GlyLulysThrAlaGlyProsnLeuAsnSerValArgGlyPheGluValLleAspThr 100  
DB 285 GGGAGAAAGACGGCCGCCCCCAACCAATTCGCGCGCTTGAGGTATCGACACC 344

QY 101 IleLysArgAsnValGluAlaAlaCysProGlyValValSerCysAlaAspLleLeuAla 120  
DB 345 ATCAAGACCAACGTGAGGCGCGCTGCCCGGCTGTGCGTGGCGGACATCTTGCC 404

QY 121 LeuAlaAlaArgAspGlyThrAsnLeuLeuGlnGlyLysProThrTrpSerValProLeuGly 140  
DB 405 CTTCCTCCGCGAGACGAGACCAATCTGCTGCGGCGCGACCTGGAACGTGCGCTGGGC 464

QY 141 ArgArgAspSerThrThrAlaSerAlaSerLleuAlaAsnSerAsnProProProThr 160  
DB 465 CGGGGGAGCTGACAGAGCGGAGCGCTCCCTTCCCAACGACCACTCCCGCAGTCGACG 524

QY 161 AlaSerLeuGlyThrLeuLleSerLeuPheGlyArg 172  
DB 525 GCGAGCCTGGGTACGCTCATCTGCTGTGACGAGG 560

RESULT 15  
LOCUS BE358471 570 bp mRNA linear EST 20-JUL-2000  
DEFINITION DGI\_30\_A01\_b1\_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA  
ACCESSION BE358471  
VERSION BE358471.1 GI:9300028  
KEYWORDS EST.  
SOURCE sorghum.  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.  
REFERENCE 1 (bases 1 to 570)  
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt,  
L.H.

TITLE An EST database from Sorghum: dark-grown seedlings  
JOURNAL Unpublished (2000)  
COMMENT Contact: Cordonier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.

Seq primer: JEN REV  
High quality sequence stop: 566  
POLYA-No.

FEATURES  
source location/Qualifiers

1..570  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Dark Grown 1 (DG1)"  
/note="Organ: 5-day-old dark-grown seedlings; Vector:  
lambda Zap; Site.1: XhoI; Site.2: EcoRI; The library was  
made from poly-A RNA in the cloning vector lambda Zap II.  
Clones to be sequenced were prepared by mass excision."  
BASE COUNT 95 a 198 c 167 g 110 t  
ORIGIN

## Alignment Scores:

Pred. No.:	9.88e-72	Length:	570
Score:	793.00	Matches:	155
Percent Similarity:	94.19%	Conservative:	7
Best Local Similarity:	90.12%	Mismatches:	10
Query Match:	49.44%	Indels:	0
DB:	10	Gaps:	0

US-10-047-825-4 (1-313) x BE358471 (1-570)

```
Oy 1 MetAlaSerProThrLeuMetGlnCysLeuValAlaValSerLeuLeuSerCysValAla 20
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 55 ATGGCGGCTCTACCTGATGCATGCTGGCGCCATTCCCTCCTCGTGTGCGC 114

Oy 21 HisAlaGlnLeuSerProThrPheThrAlaSerSerCysProAsnLeuGlnSerIleVal 40
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 115 CAGGCACAGCTCTGCAGCCGTTCTATGATCTCTCTGCCCACTGCAGACATCGTG 174

Oy 41 ArgAlaAlaMetThrGlnAlaValAlaSerGlnGlnArgMetGlyAlaSerLeuLeuArg 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 175 CGGAGGCGCATGATCAGGCCCTCAGCAATGACAGAGATGGCGCTCTGTGCTCAGG 234

Oy 61 LeuPhePheHisAspCysPheValGlnGlyCysAspGlySerIleLeuLeuAspAlaGly 80
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 235 CTCTCTTCACAGACTGCTTGTCAAGTTCGACGGCTCGATCCTCTCGACGCCGGA 294

Oy 81 GlyGlnGlyThrAlaGlyProAsnLeuAsnSerValArgGlyPheGluValIleAspThr 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 295 GGGGAGAGAGAGCGGCCGCCCAACGCCAATTCGGCGCGGCTTGTGAGTCAATCGACACC 354

Oy 101 IleLysArgAsnValGlnAlaIaIaCysProGlyValIaIaSerCysAlaAspIleLeuAla 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 355 ATCAACACCAACAGCTGAGGCCCGGTGCCGCTGTGCTGTGCGGACATCCTTGCC 414

Oy 121 LeuAlaIaIaArgAspGlyThrAsnLeuLeuGlyGlyProThrTrpSerValProLeuGly 140
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 415 CTTCGCGCGGAGACGACCAATCTGCTCGCGGCGGACCTGGAACGTGCGCGCTGGGC 474

Oy 141 ArgArgAspSerThrThrAlaSerAlaSerLeuAlaAsnSerAsnProProProProThr 160
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 475 CGCGGGGACTCGACGCGGCGGAGCGGCTCTGTGCCAACAGACCTCCGCGAGTCGACG 534

Oy 161 AlaSerLeuGlyThrLeuIleSerLeuPheGlyArg 172
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 535 GCGAGCGCTGGTACGCTCATCTCGCTGTTCAGCAGG 570
```

Search completed: May 3, 2003, 13:15:25  
Job time: 1089 secs